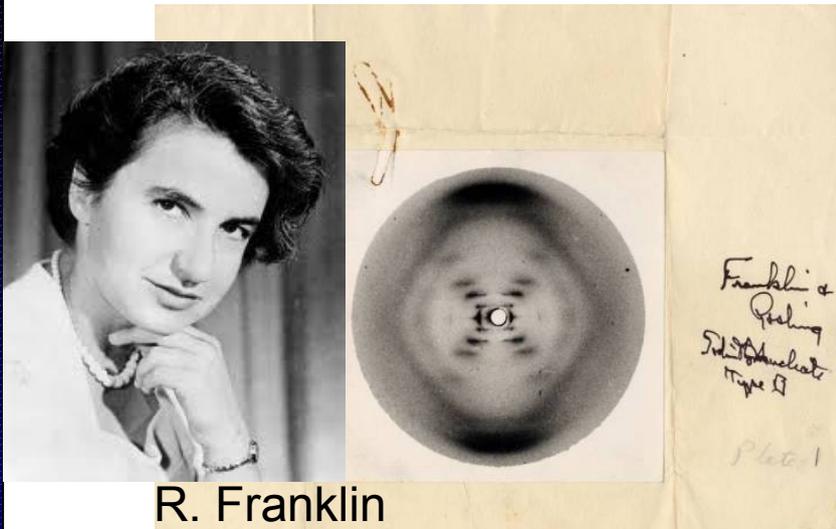


Épigénétique et variabilités génomiques

Journée de formation
Évolution des génomes, évolution de l'Homme.
14 Mai 2013

thomas.heams@agroparistech.fr



R. Franklin

14 mai 2013

Épigénétique et Variabilités Génomiques

F. Crick & J. Watson

April 25, 1953 NATURE

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribonucleic Acid

WE wish to suggest a structure for the salt of deoxyribonucleic acid (D.N.A.). The structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey¹. They briefly made their manuscript available to us in advance of publication. Their model consists of three inter-linked chains, with the phosphates near the three ends and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the various chains given the X-ray diagram in the salt, run the free ends. Without the various hydrogens shown it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the ends will repel each other. (2) Some of the sites for which distances appear to be too small.

Another alternative structure has also been suggested by Pauling in the press.² In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for his reasons we shall not comment on it.

We wish to put forward a radically different structure for the salt of deoxyribonucleic acid. This structure has two linked chains each coiled round the same axis (see diagram). We have made the usual standard assumptions, namely, that each chain consists of phosphate alternate groups joined to 2-deoxy-ribose with N.P. linkages. The two chains do not their bases are joined by a bond perpendicular to the three axis. Both chains follow right-handed helices, but owing to the spiral the sequence of the bases in the two chains run in opposite directions. Each chain closely resembles Pauling's model No. 1,¹ that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the groups near it is close to Pauling's "preferred configuration", the sugar being roughly perpendicular to the three-axis line. There is a rotation on each chain every 3.4 Å. in the z-direction. We have assumed an angle of 36° between adjacent bases in the same chain, so that the sequence repeats after 10 rotations on each chain, that is, after 34 Å. The distance of a phosphate group from the three axis is 10 Å. All the phosphates are on the outside, various base pairs occur in turn.

This structure is in open form, and its water content is rather high. As liquid water contents are usually expected the bases are able to form the structure could become more compact.

The special feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the three axis. They are joined

together by hydrogen bonds, two towards the ribe hydrogens and two towards the phosphate groups. It is a structure that is appropriate to a typical "protein" (protein). In such a pair, the other groups are simple enough to say. It follows, in future work if the sequence of bases on one chain is given, that the sequence on the other chain is automatically determined.

It has been found experimentally^{3,4} that the ratio of guanine to cytosine, and always very close to unity, the deoxyribose nucleic acid.

It is probably impossible to build this structure with a ribbon model in place of the deoxyribose, as the extra oxygen atoms would make too close a run for the three axis.

The previously published X-ray data⁵ on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. As far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in the following communications. We have not shown all the details of the model presented there where we do not have data of sufficient accuracy to justify our strictly on published experimental data and theoretical requirements.

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

Full details of the structure, including the distances assumed in building it, together with a list of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on inter-atomic distances. We have also been privileged by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. M. H. P. Williams, Dr. H. K. Franklin and their co-workers at King's College, London. One of us (J. D. W.) has been aided by a fellowship from the National Foundation for Infectious Diseases.

J. D. WATSON
F. H. C. CRICK
Medical Research Council Unit for the Study of the Molecular Structure of Biological Systems,
Cavendish Laboratory, Cambridge.
(April 2)

¹ Pauling, L., and Corey, R. B., *Proc. Nat. Acad. Sci. U.S.A.*, **37**, 252 (1951).
² Pauling, L., and Corey, R. B., *ibid.*, **37**, 252 (1951).
³ Chargaff, E., *Proc. Nat. Acad. Sci. U.S.A.*, **36**, 258 (1950).
⁴ Chargaff, E., *Proc. Nat. Acad. Sci. U.S.A.*, **36**, 258 (1950).
⁵ Franklin, R. E., *Proc. Nat. Acad. Sci. U.S.A.*, **37**, 35 (1951).
⁶ Franklin, R. E., *Proc. Nat. Acad. Sci. U.S.A.*, **37**, 35 (1951).

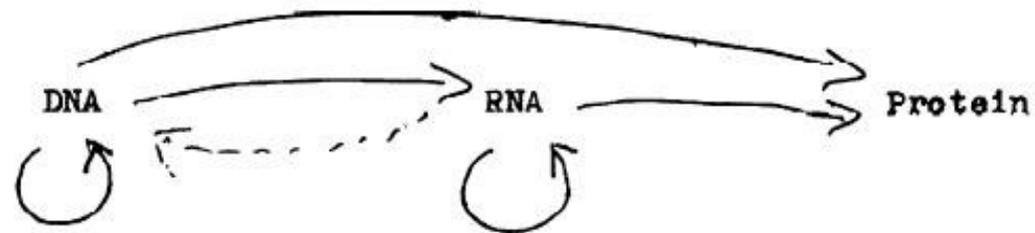


Ideas on Protein Synthesis (Oct. 1956)

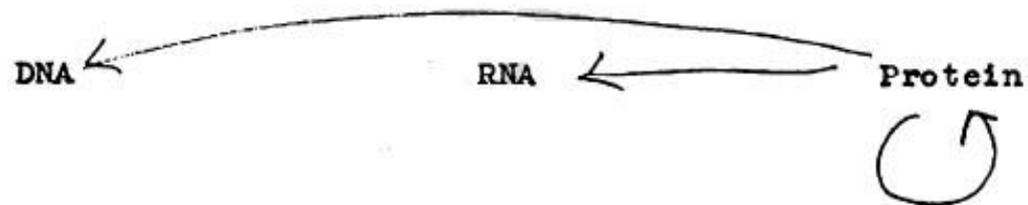
The Doctrine of the Triad.

The Central Dogma: "Once information has got into a protein it can't get out again". Information here means the sequence of the amino acid residues, or other sequences related to it.

That is, we may be able to have



but never



where the arrows show the transfer of information.



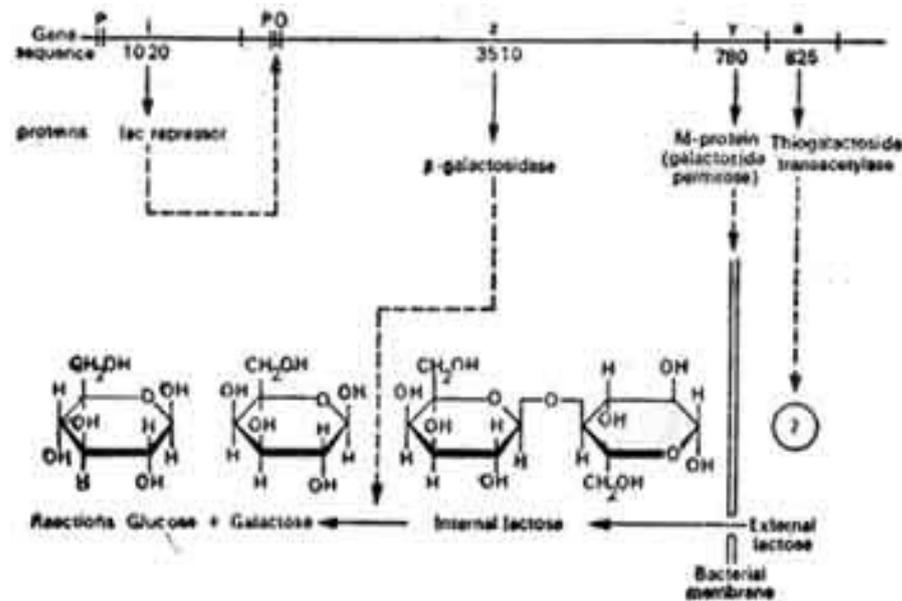
REVIEW ARTICLE

Genetic Regulatory Mechanisms in the Synthesis of Proteins †

FRANÇOIS JACOB AND JACQUES MONOD

*Services de Génétique Microbienne et de Biochimie Cellulaire,
Institut Pasteur, Paris*

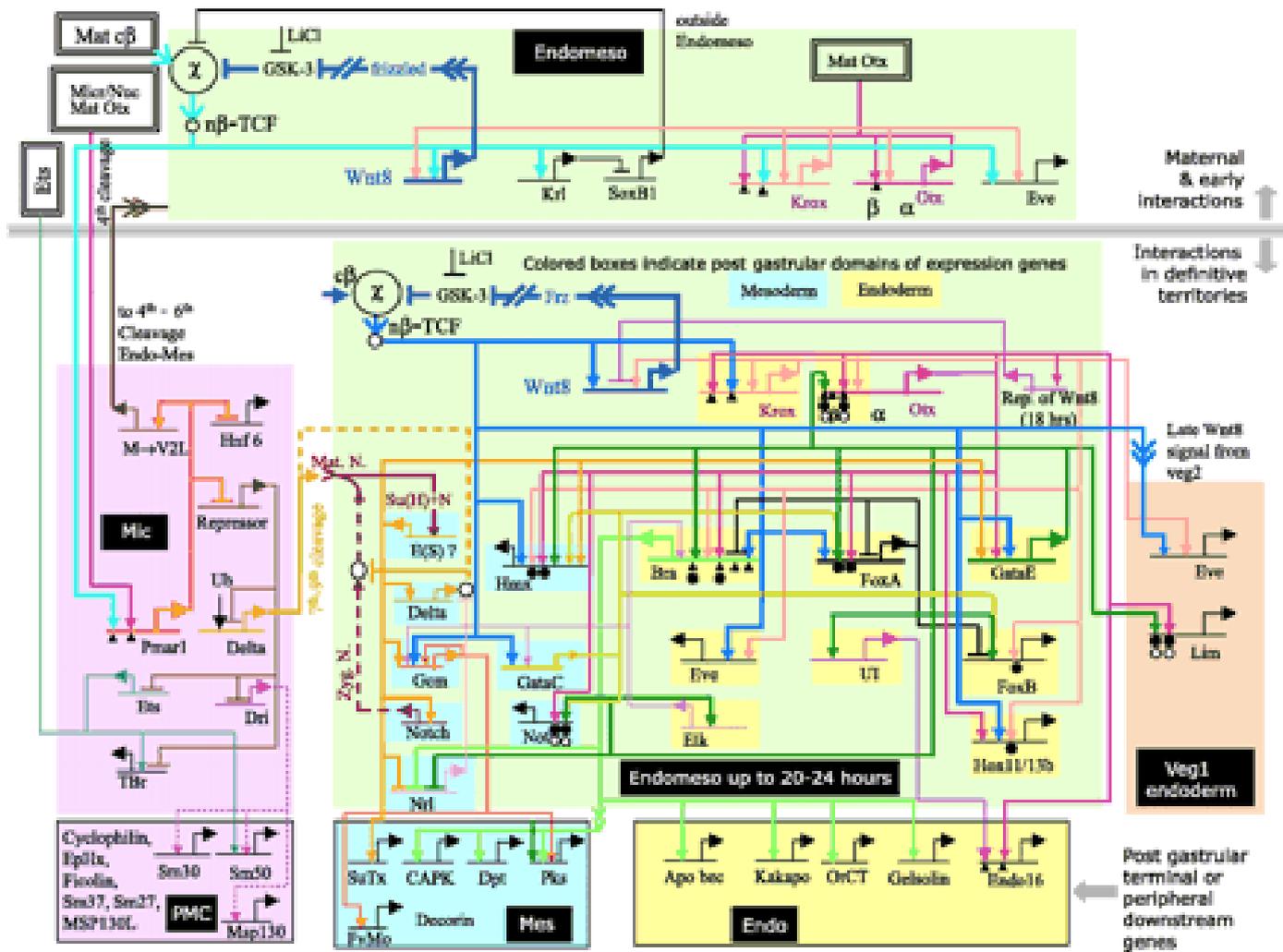
(Received 28 December 1960)



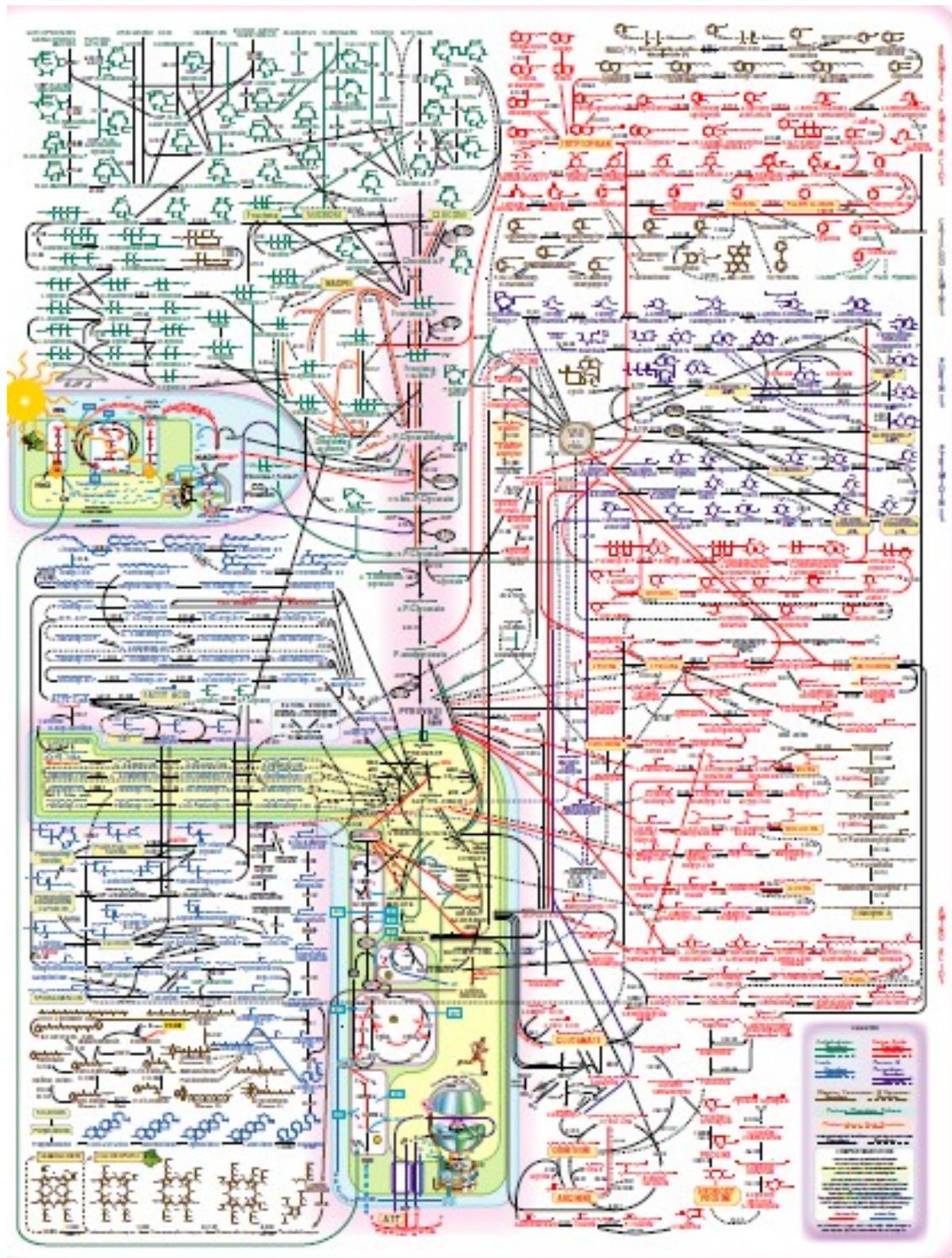
14 mai 2013

Épigénétique et Variabilités Génomiques

« A Genomic Regulatory Network for Development »



Davidson *et al*, 2002



© Sigma-Aldrich, 2003

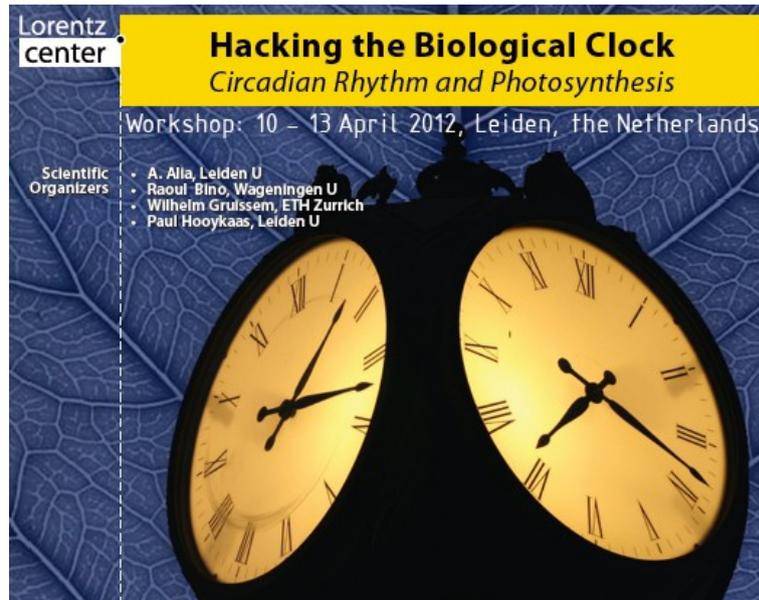
LETTER

doi:10.1038/nature11149

Programmable single-cell mammalian biocomputers

Simon Ausländer¹, David Ausländer¹, Marius Müller¹, Markus Wieland¹ & Martin Fussenegger^{1,2}

Nature 2010



REVIEW ARTICLE

Bacteria as computers making computers

Antoine Danchin

Génétique des Génomes Bactériens, Institut Pasteur, Paris, France

FEMS Microbiol Rev 2009

Comparing genomes to computer operating systems in terms of the topology and evolution of their regulatory control networks

Koon-Kiu Yan^a, Gang Fang^a, Nitin Bhardwaj^a, Roger P. Alexander^a, and Mark Gerstein^{b,a,c,1}

PNAS 2010

14 mai 2013

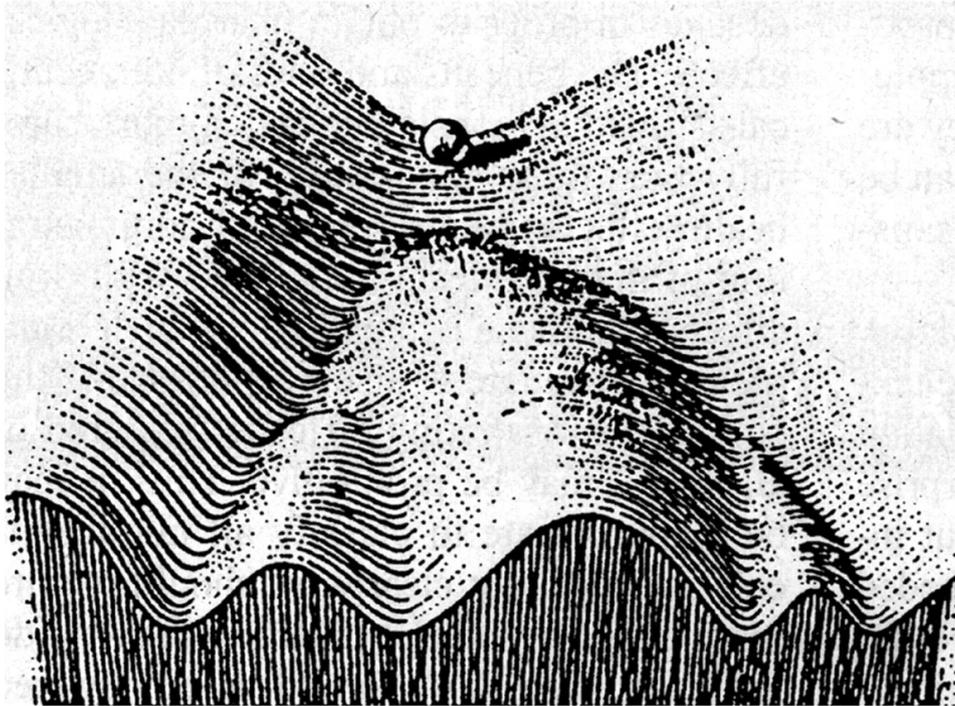
Épigénétique et Variabilités Génomiques

Tout invite à assimiler la logique de l'hérédité à celle d'une calculatrice.
Rarement modèle imposé par une époque aura trouvé une application plus fidèle.



François Jacob,
la Logique du Vivant, 1970

Epigénétique

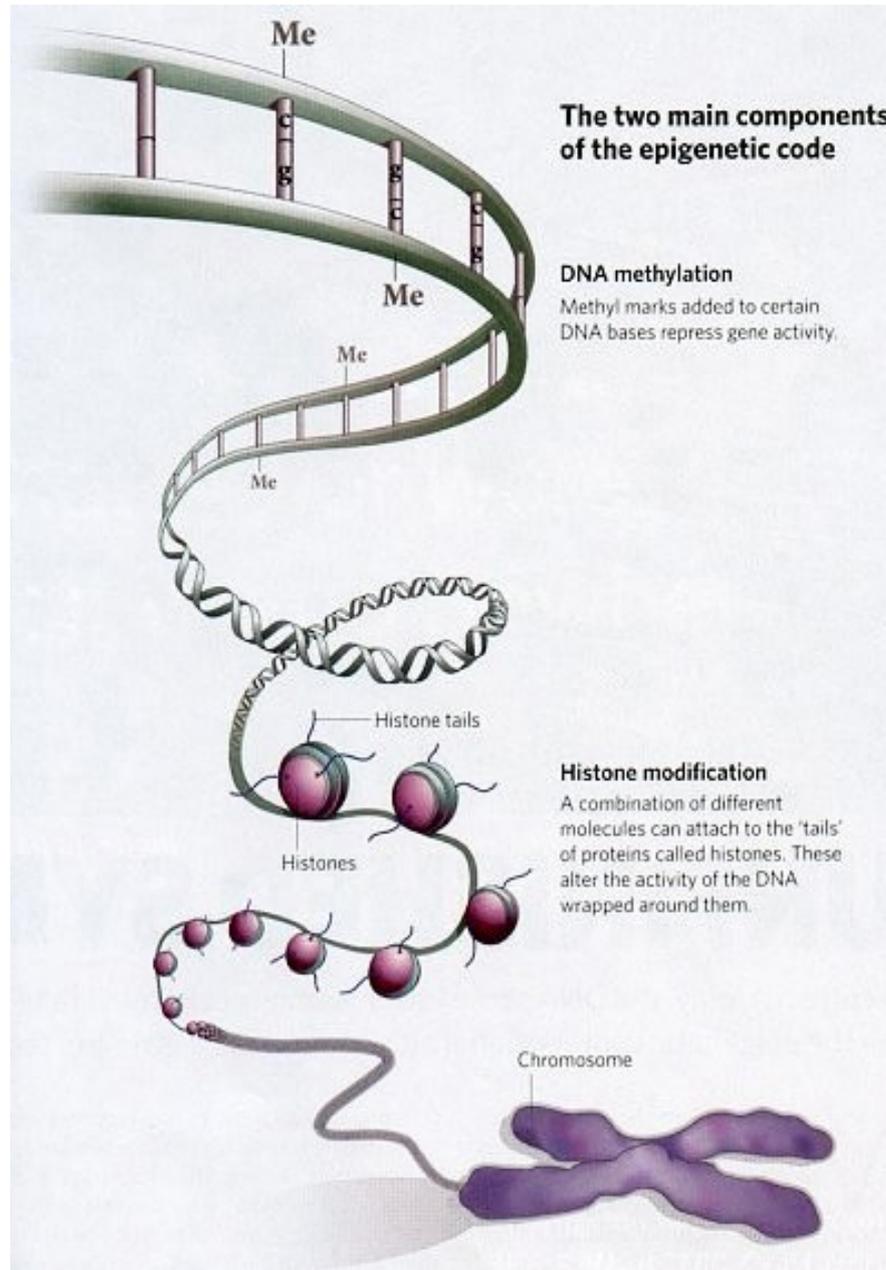


Conrad Waddington
« *Epigenetics* » 1942

Paysage épigénétique (tiré de *Strategy of the genes*, 1957)

L'épigénétique Une nouvelle frontière (?)

Modifications
-fonctionnelles à séquence constante
-(partiellement) héritables



Nature 441: 143-145. 11 May 2006

Modifications
chimiques
Sur l'ADN

Modification dans l'organisation
3D de la chromatine
=Territoires chromosomiques

Interférence ARN

Expression
aléatoire

14 mai 2013

Epigénétique et Variabilités Génomiques

Méthylation de l'ADN

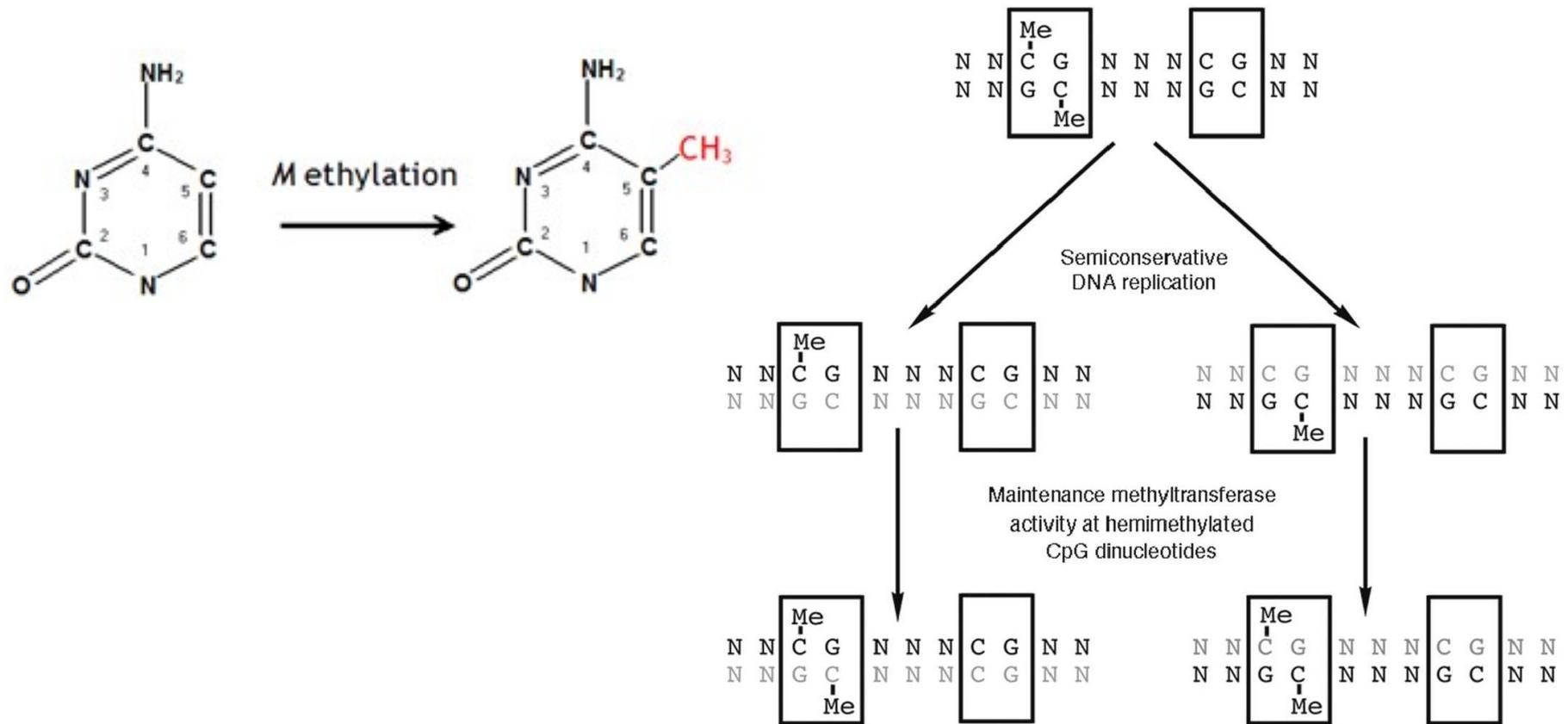


Figure 1. Maintenance methylation at CpG sites in mammalian DNA. The current view of DNA methylation in mammals involves de novo methylation in germ cells of the previous generation and in the early embryo, followed by maintenance methylation during development. Hemimethylated sites are thought to be the target for DNA methyltransferases of the DNA methyltransferase-1 (Dnmt1) family. As discussed in the text, this model is probably oversimplified to some degree.

<http://www.bio.miami.edu>

DNA Replication and Human Disease © 2006 Cold Spring Harbor Laboratory Press, Chapter 7, Figure 1.

14 mai 2013

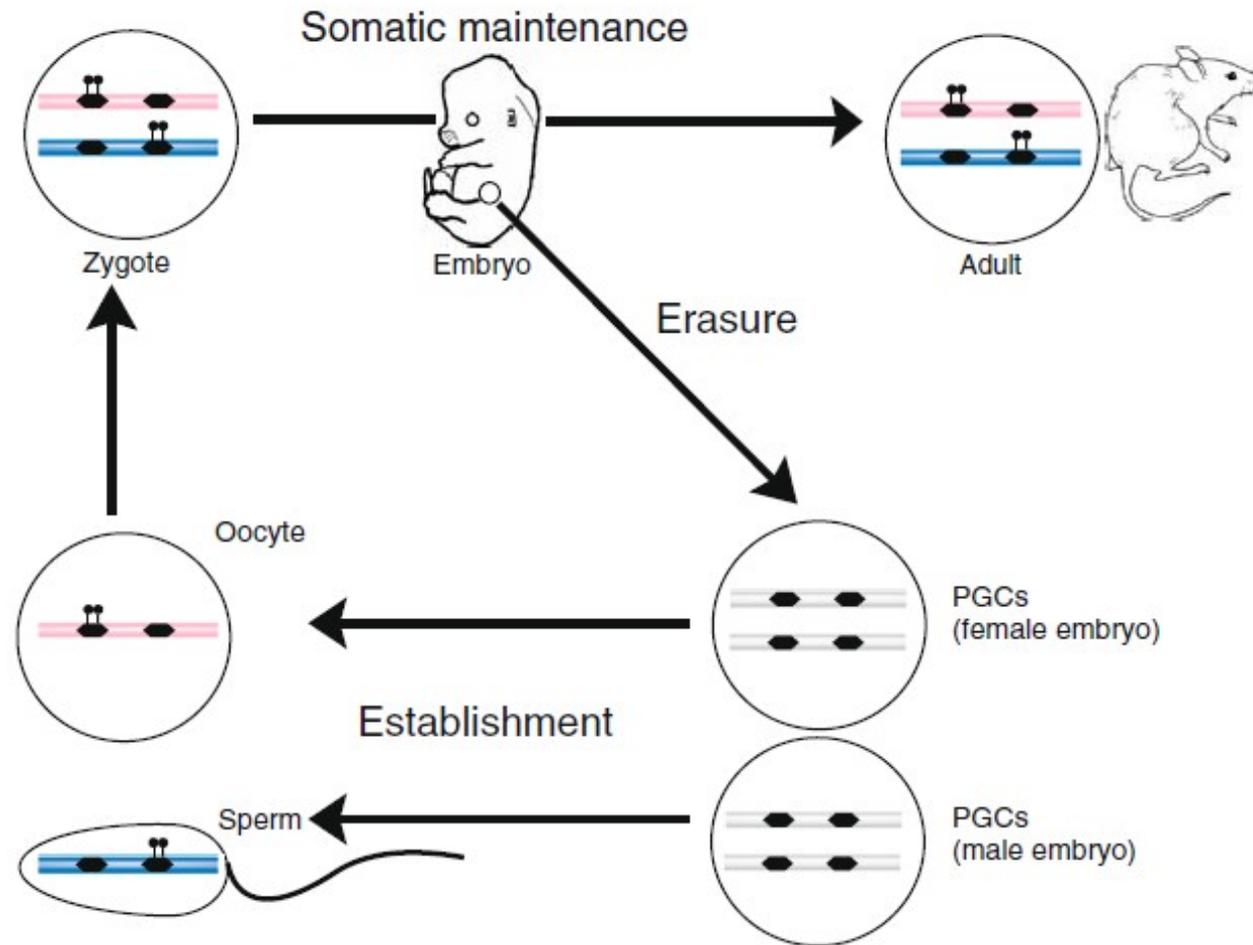
Épigénétique et Variabilités Génomiques

La méthylation peut être transmise de parent à enfant dans la lignée germinale, et est centrale dans certaines formes d'hérédité épigénétique

La méthylation est **essentielle au développement normal de l'embryon** : elle permet d'éteindre certains gènes qui n'ont plus à être exprimés une fois qu'ils ont joué leur rôle, ou d'atténuer l'expression génétique pour maintenir l'identité cellulaire.

La méthylation de l'ADN **pourrait avoir évolué comme un mécanisme de défense** contre l'ADN viral, qui est inactivé par des enzymes de méthylation de l'ADN. L'existence de fragments d'ADN méthylés, silencieux, tout au long de notre génome pourrait être la trace d'anciens virus parasites dégradés et séquestrés

Empreinte génomique : méthylation allélique différentielle



Kacem & Feil, 2009

Exemple : Igf2 chez la souris (Insulin-like growth factor gene)

Un des premiers gènes soumis à empreinte découvert est Igf2 chez la souris. Le produit de ce gène est nécessaire pour la croissance normale de l'embryon, mais seul l'allèle paternel est exprimé. La copie maternelle est très méthylée, et silencieuse

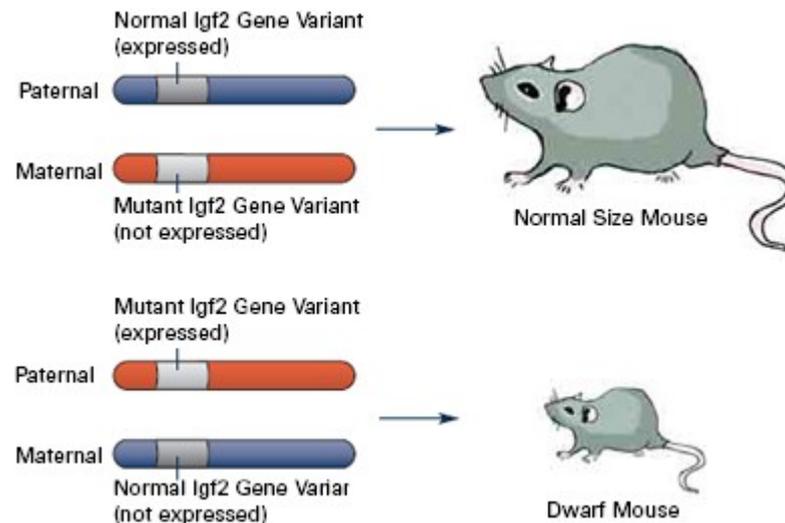
Le gène peut muter et devenir inactif dans chaque gamète:

Ovocyte : une mutation résulte en un allèle maternel non fonctionnel

Spermatozoïde: une mutation résulte en un allèle paternel non fonctionnel

Comme l'allèle maternel est silencieux, hériter d'un allèle maternel muté n'aura pas d'effet : le souriceau grossira normalement

Mais à l'inverse, hériter d'un allèle paternel muté conduira à l'expression de ce seul allèle et donc à du nanisme



<http://www.bio.miami.edu>

High-Resolution Analysis of Parent-of-Origin Allelic Expression in the Mouse Brain

2010

Christopher Gregg,^{1,2*†} Jiangwen Zhang,^{3*} Brandon Weissbourd,^{1,2} Shujun Luo,⁵ Gary P. Schroth,⁵ David Haig,⁴ Catherine Dulac^{1,2†}

Genomic imprinting results in preferential expression of the paternal or maternal allele of certain genes. We have performed a genome-wide characterization of imprinting in the mouse embryonic and adult brain. This approach uncovered parent-of-origin allelic effects of more than 1300 loci. We identified parental bias in the expression of individual genes and of specific transcript isoforms, with differences between brain regions. Many imprinted genes are expressed in neural systems associated with feeding and motivated behaviors, and parental biases preferentially target genetic pathways governing metabolism and cell adhesion. We observed a preferential maternal contribution to gene expression in the developing brain and a major paternal contribution in the adult brain. Thus, parental expression bias emerges as a major mode of epigenetic regulation in the brain.

NEWS IN FOCUS

GENETICS

Mais...



RNA studies under fire

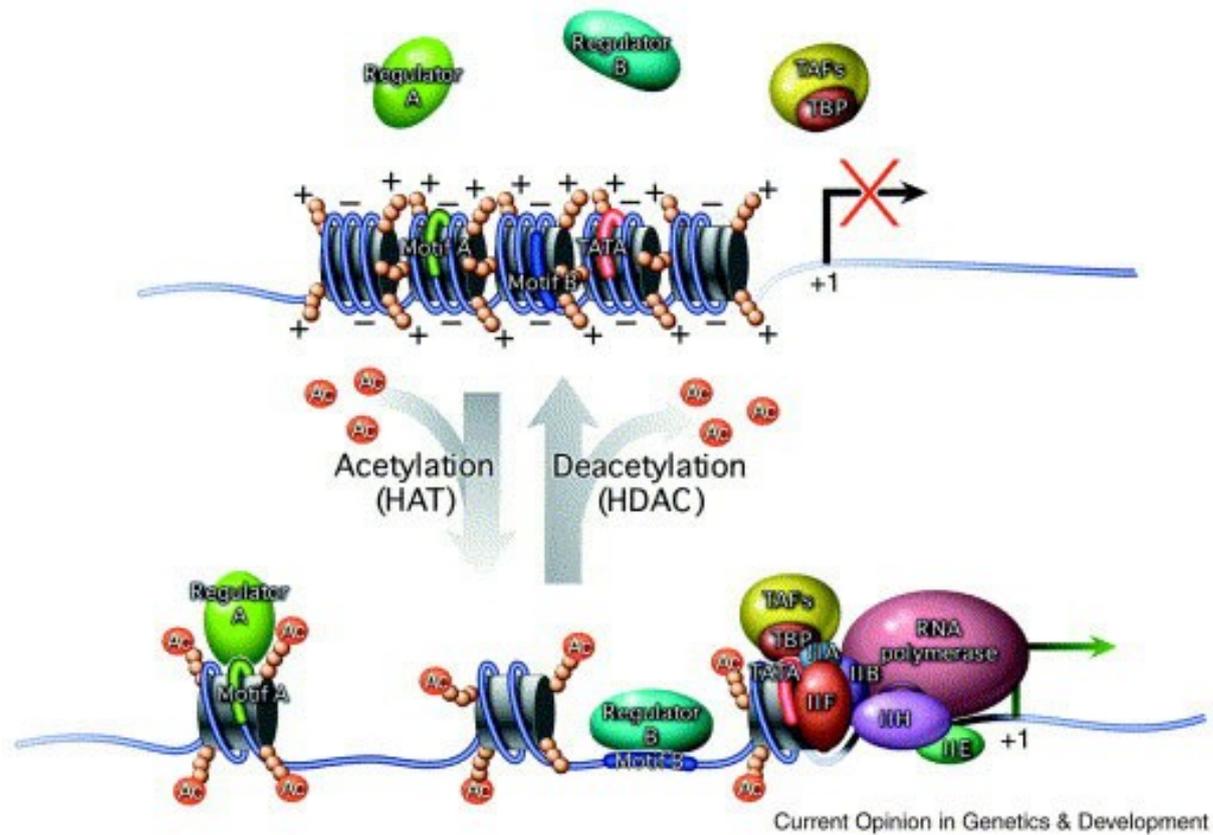
High-profile results challenged over statistical analysis of sequence data.

14 mai 2013

Épigénétique et Variabilités Génomiques

Nature News, 2012

Une autre marque épigénétique : l'acétylation des histones



D'après Hsieh & Gage, 2004

Translating the Histone Code

Thomas Jenuwein¹ and C. David Allis²

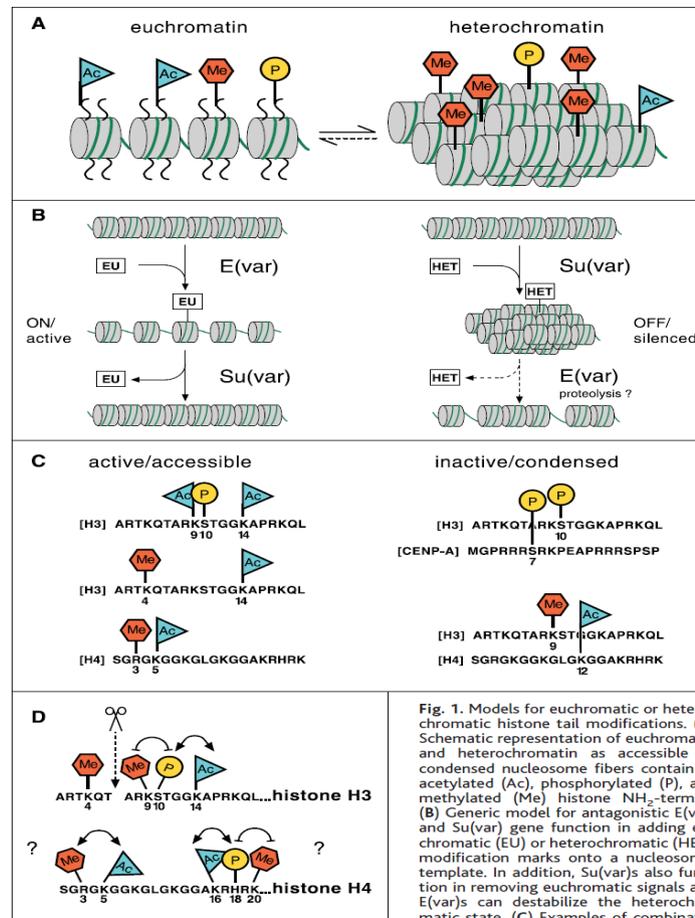
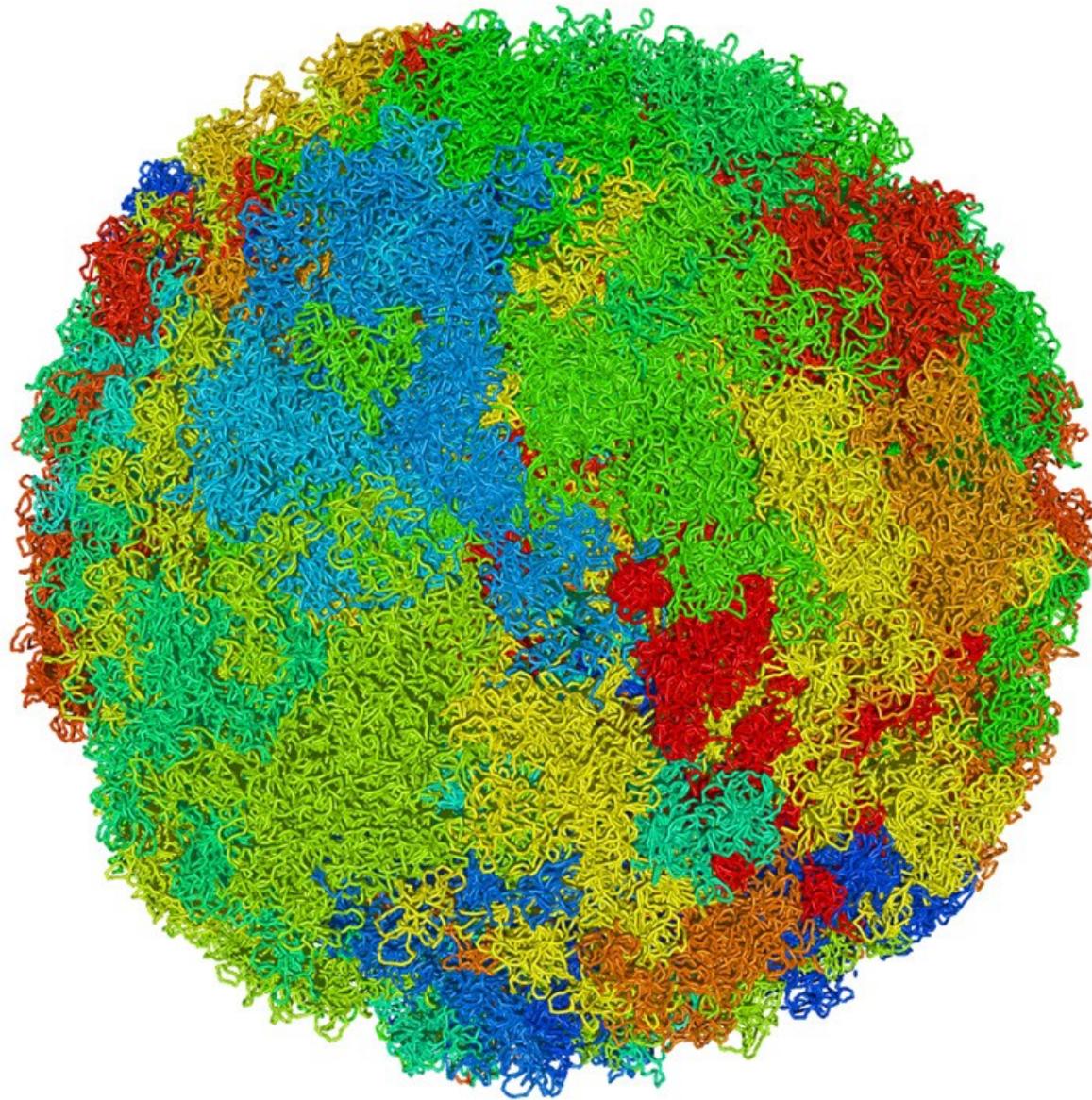


Fig. 1. Models for euchromatic or heterochromatic histone tail modifications. (A) Schematic representation of euchromatin and heterochromatin as accessible or condensed nucleosome fibers containing acetylated (Ac), phosphorylated (P), and methylated (Me) histone NH₂-termini. (B) Generic model for antagonistic E(var) and Su(var) gene function in adding euchromatic (EU) or heterochromatic (HET) modification marks onto a nucleosomal template. In addition, Su(var)s also function in removing euchromatic signals and E(var)s can destabilize the heterochromatic state. (C) Examples of combinatorial modifications in histone NH₂-termini that are likely to represent "imprints" for active or inactive chromatin. Single-letter abbreviations for amino acid residues: A, Ala; E, Glu; G, Gly; H, His; K, Lys; L, Leu; M, Met; P, Pro; Q, Gln; R, Arg; S, Ser; and T, Thr. (D) Proposed synergistic (connected arrowheads) or antagonistic (blocked oval line) modifications in histone H3 and H4 NH₂-termini. The arrow with the scissors indicates possible proteolytic cleavage of the H3 NH₂-terminus.

2001

Territoires chromosomiques



Knoch, 2002

14 mai 2013

Epigénétique et Variabilités Génomiques

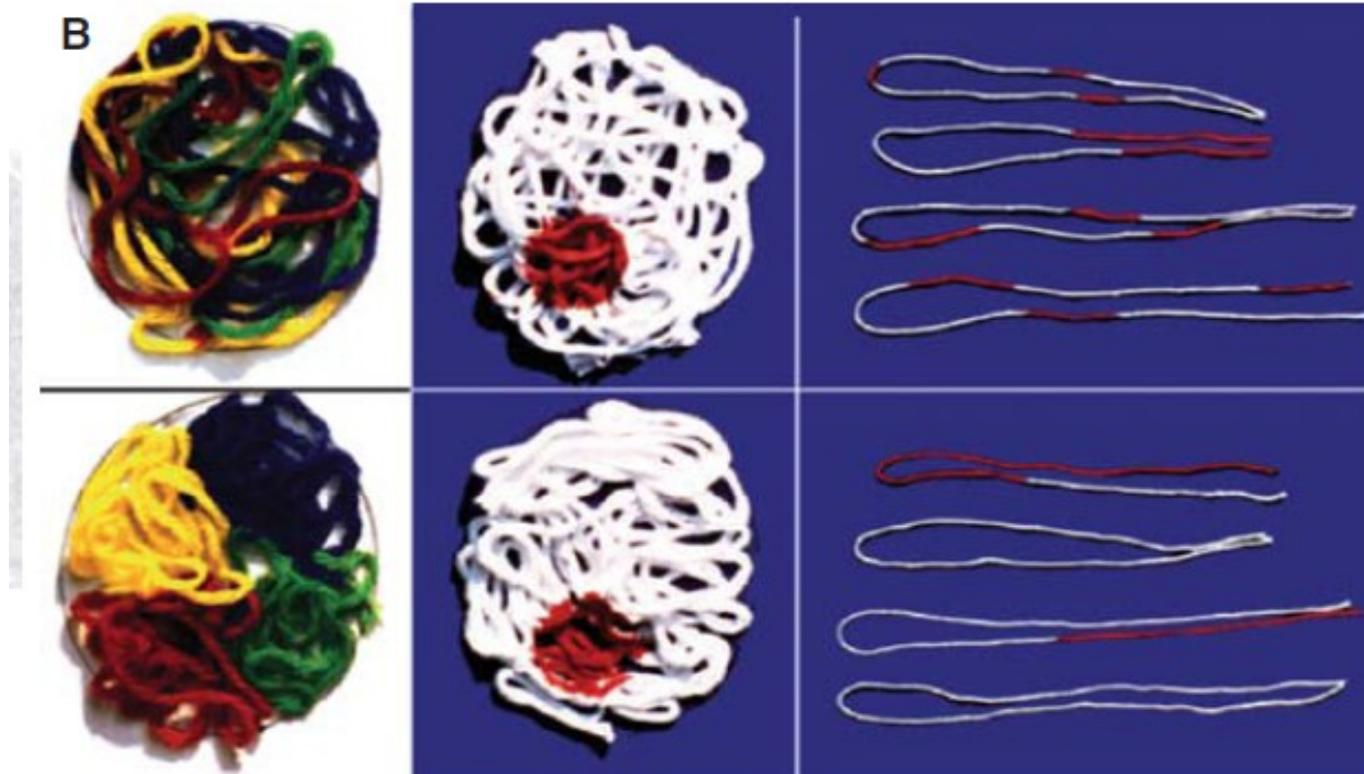


?

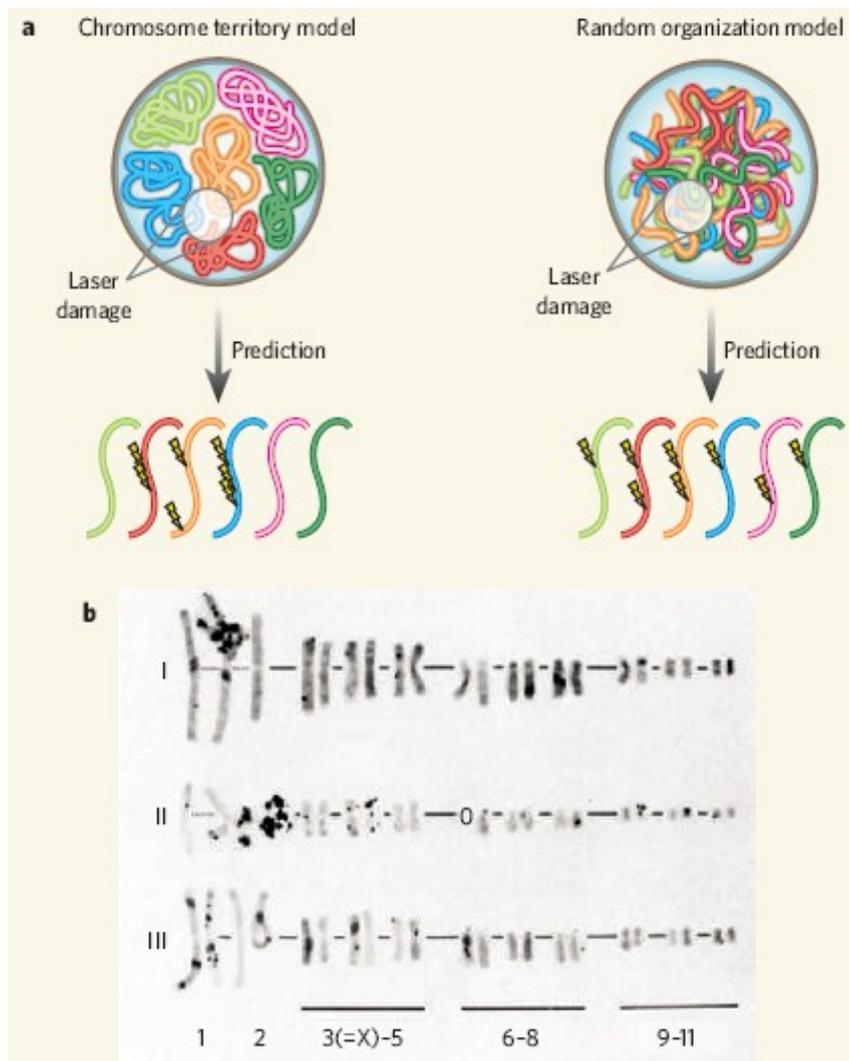
14 mai 2013

Epigénétique et Variabilités Génomiques

A la fin du 19ème siècle, C. Rabl et Th. Boveri émettent l'hypothèse que pendant l'interphase, les chromosomes gardent leur intégrité

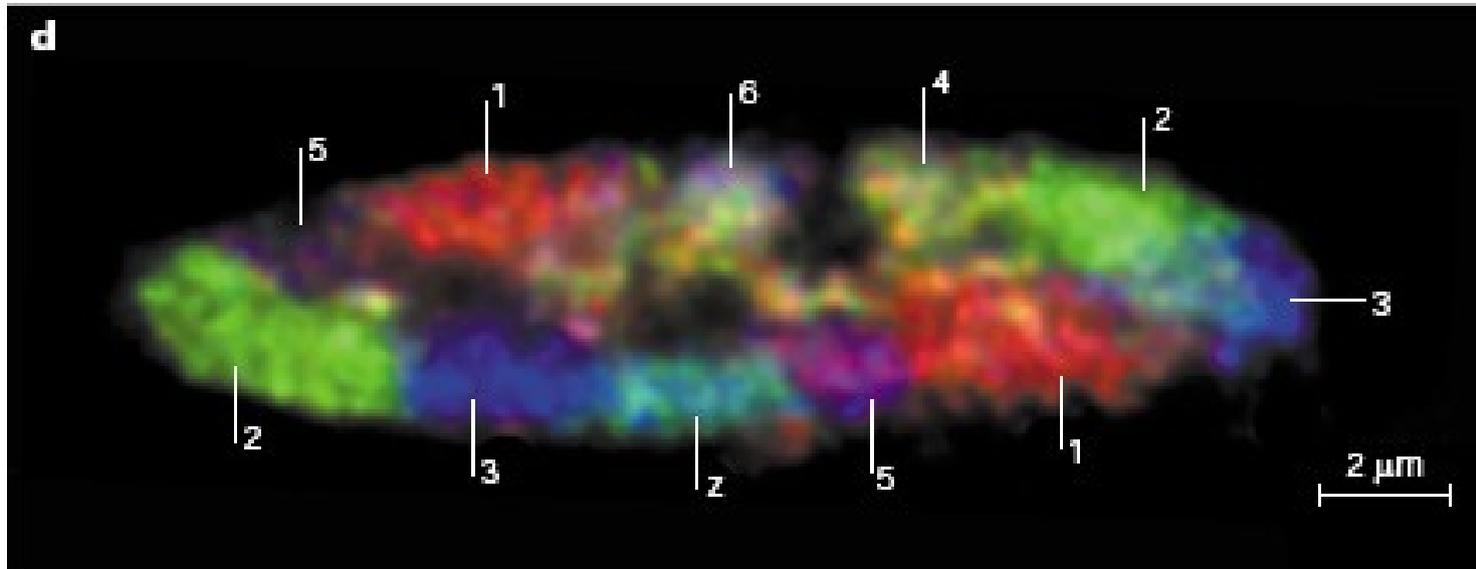


Cremer *et al*, 1982 (in Cremer & Cremer 2010)



Cremer & Cremer 2010

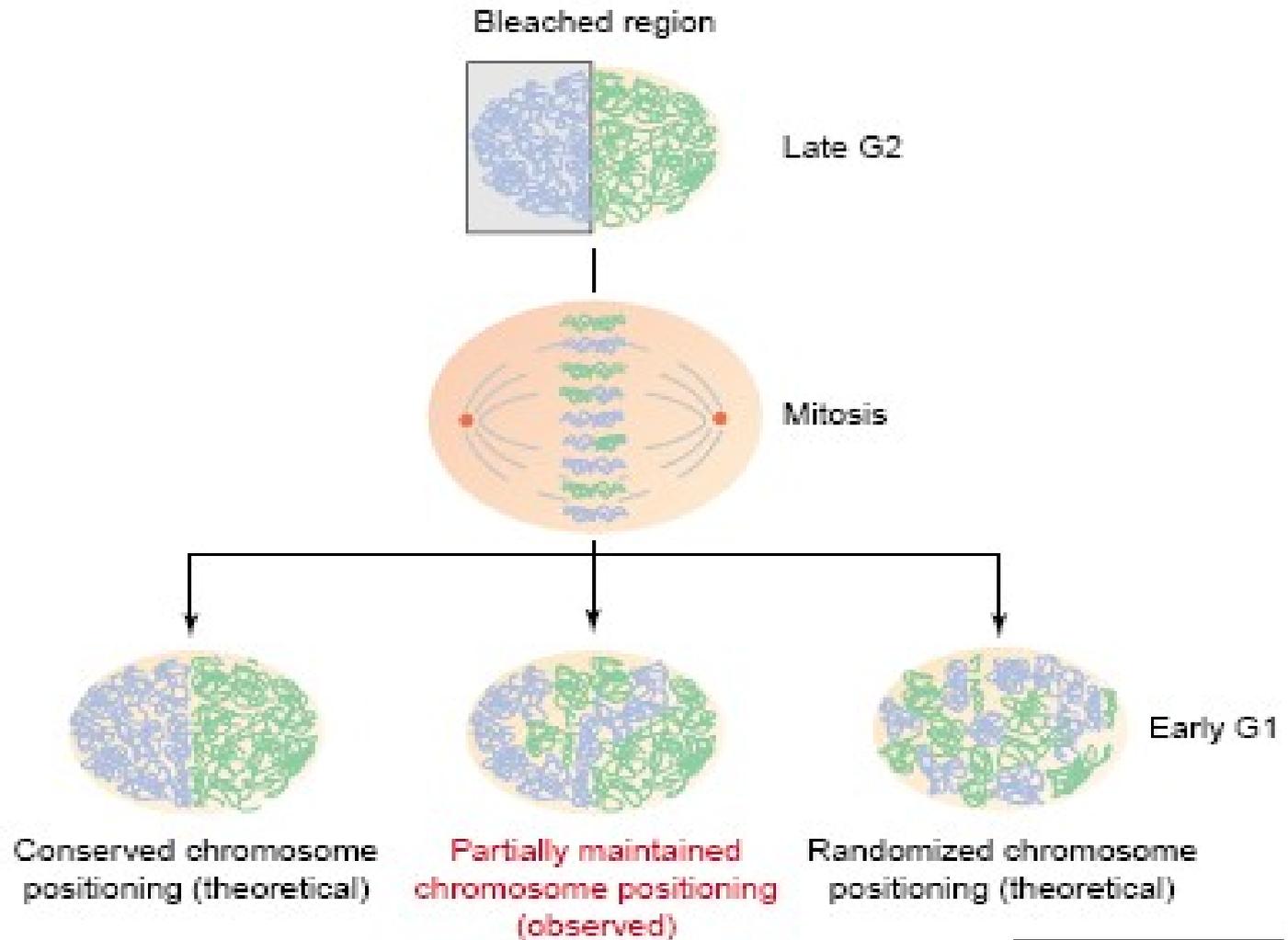
Confirmation dans les années 1980
par Thomas et Christoph Cremer



Cremer & Cremer, 2001

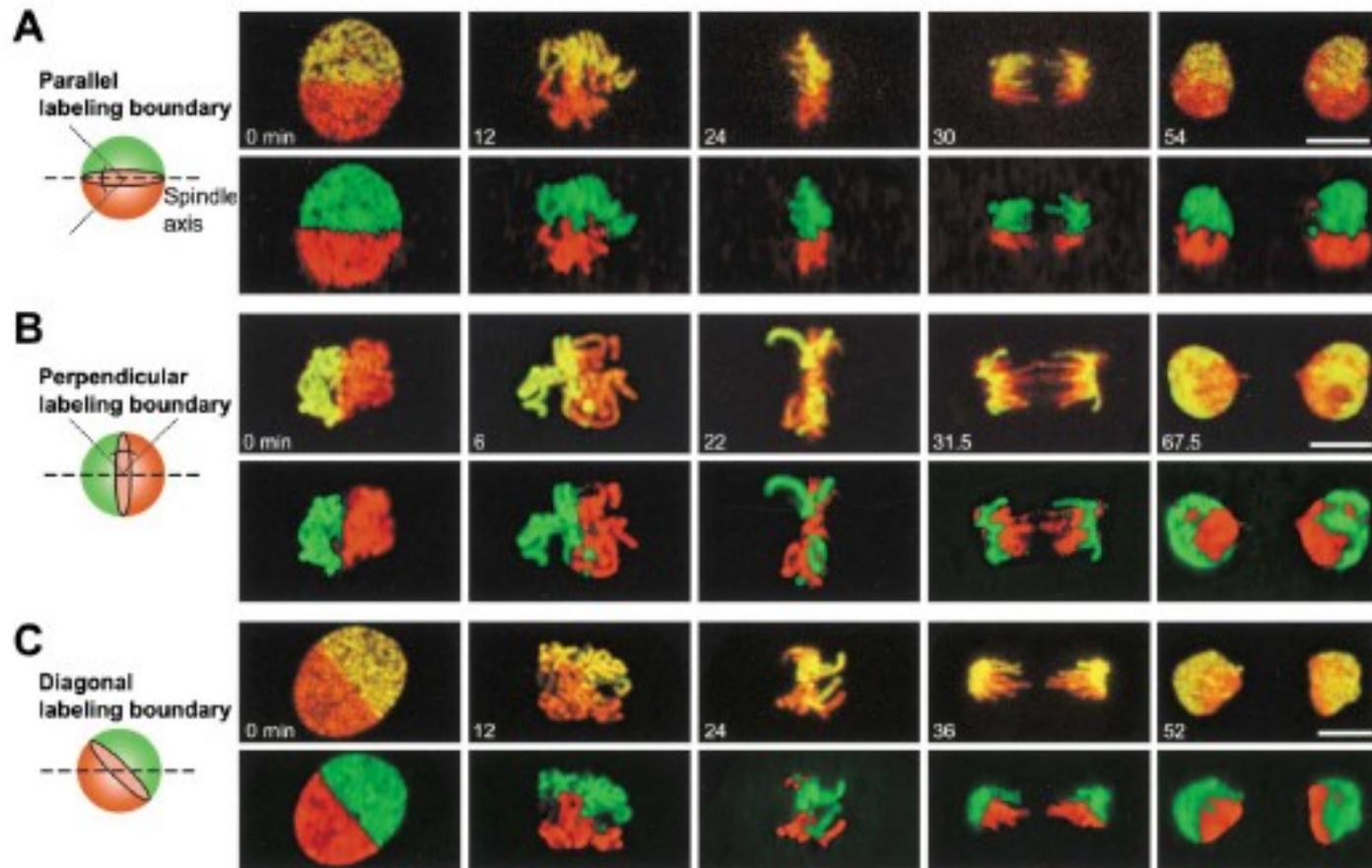
Noyau de cellule (Poulet) à l'interphase (*G. gallus*)

Mise en évidence d'une persistance partielle de la distribution interphasique après la division cellulaire



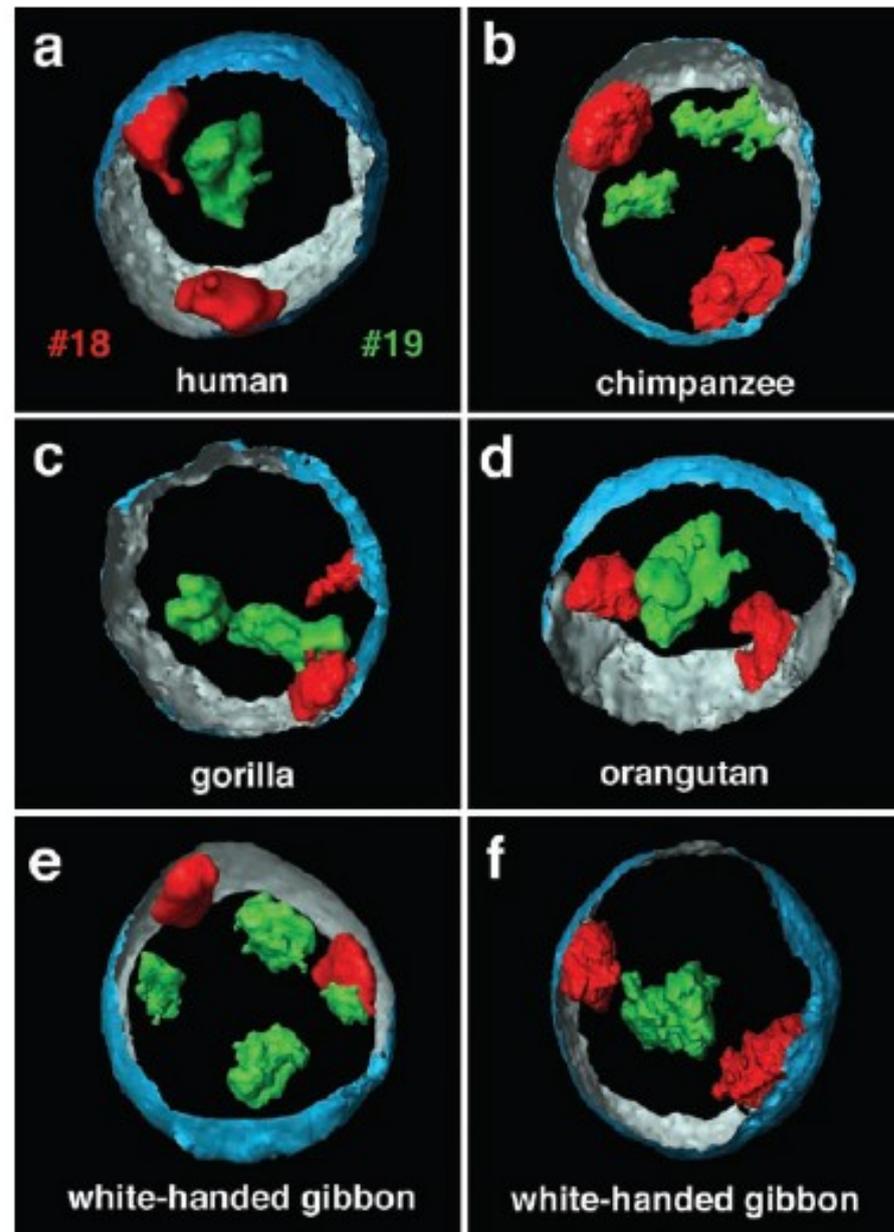
TRENDS in Cell Biology

Parada *et al*, 2003

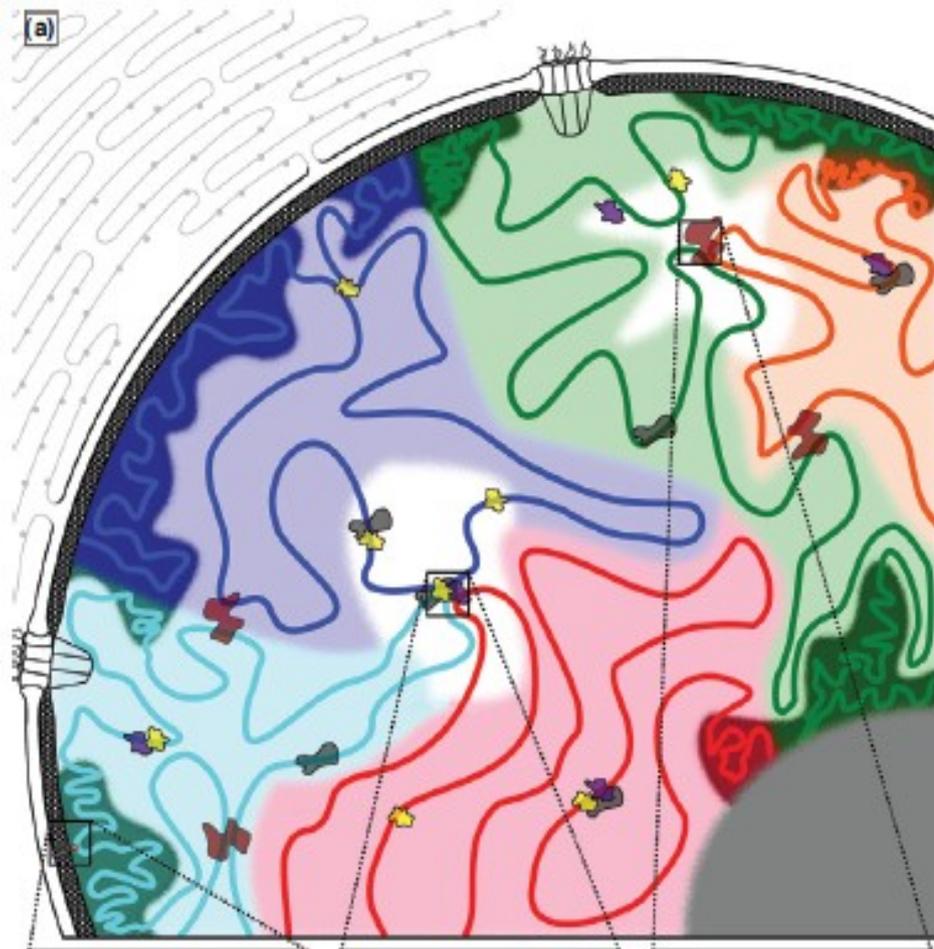


Gerlich *et al*, 2003

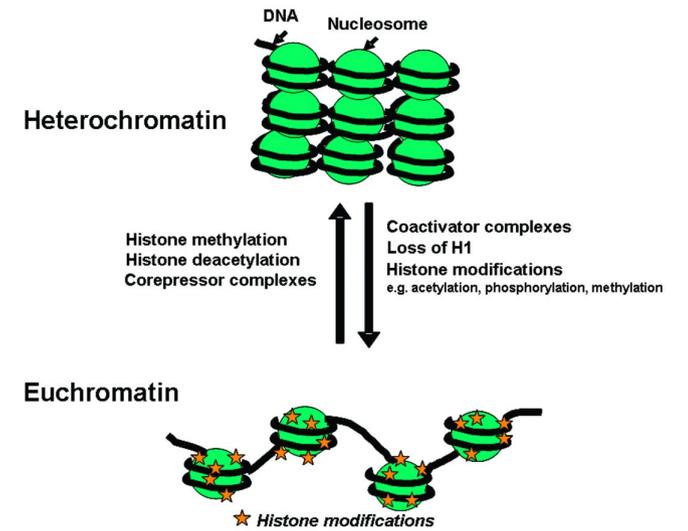
Conservation évolutive des territoires chromosomiques



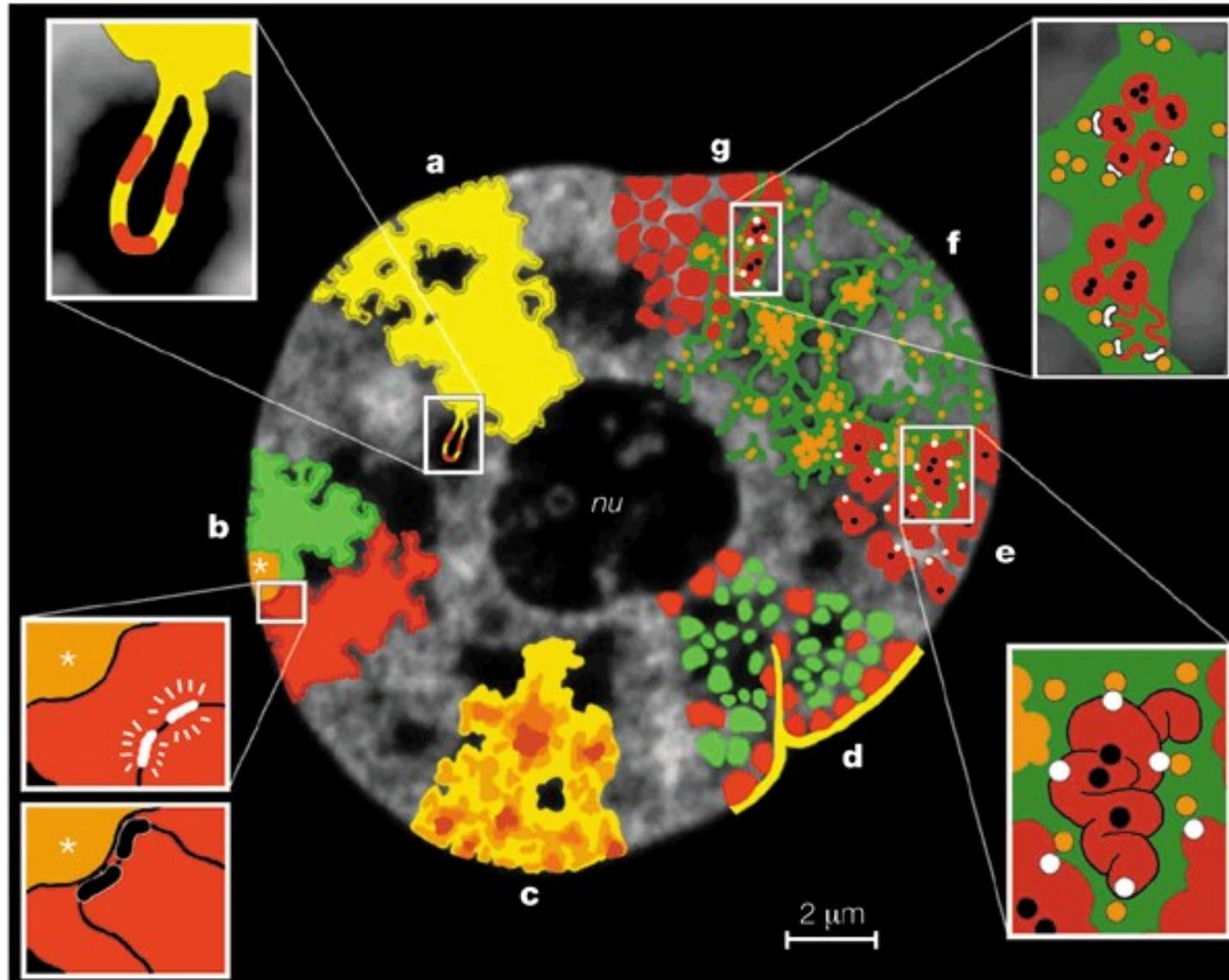
Tanabe *et al*, 2007



Geyer *et al*, 2011



Plusieurs niveaux de complexité

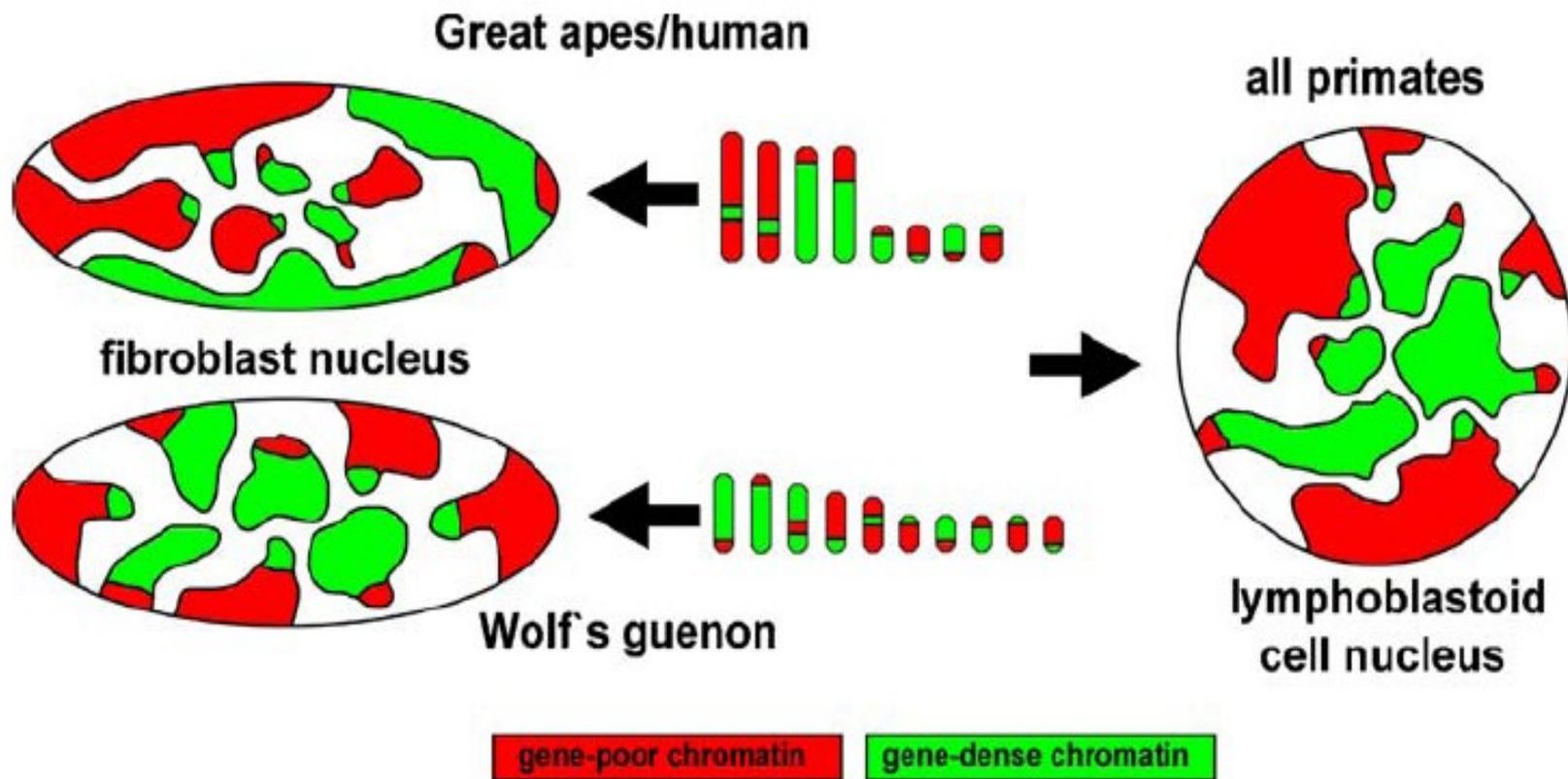


Nature Reviews | Genetics

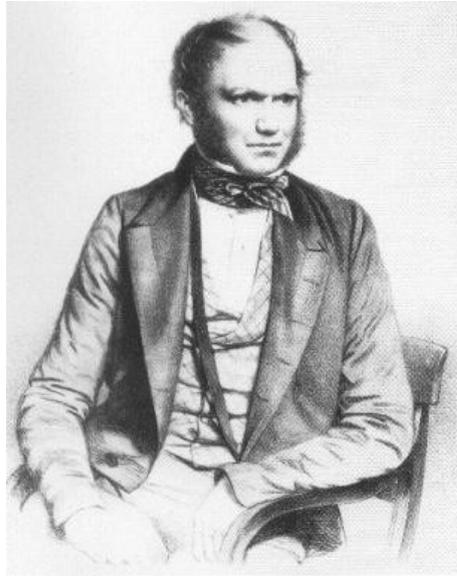
Cremer & Cremer, 2001

14 mai 2013

Epigénétique et Variabilités Génomiques



Neusser *et al*, 2007

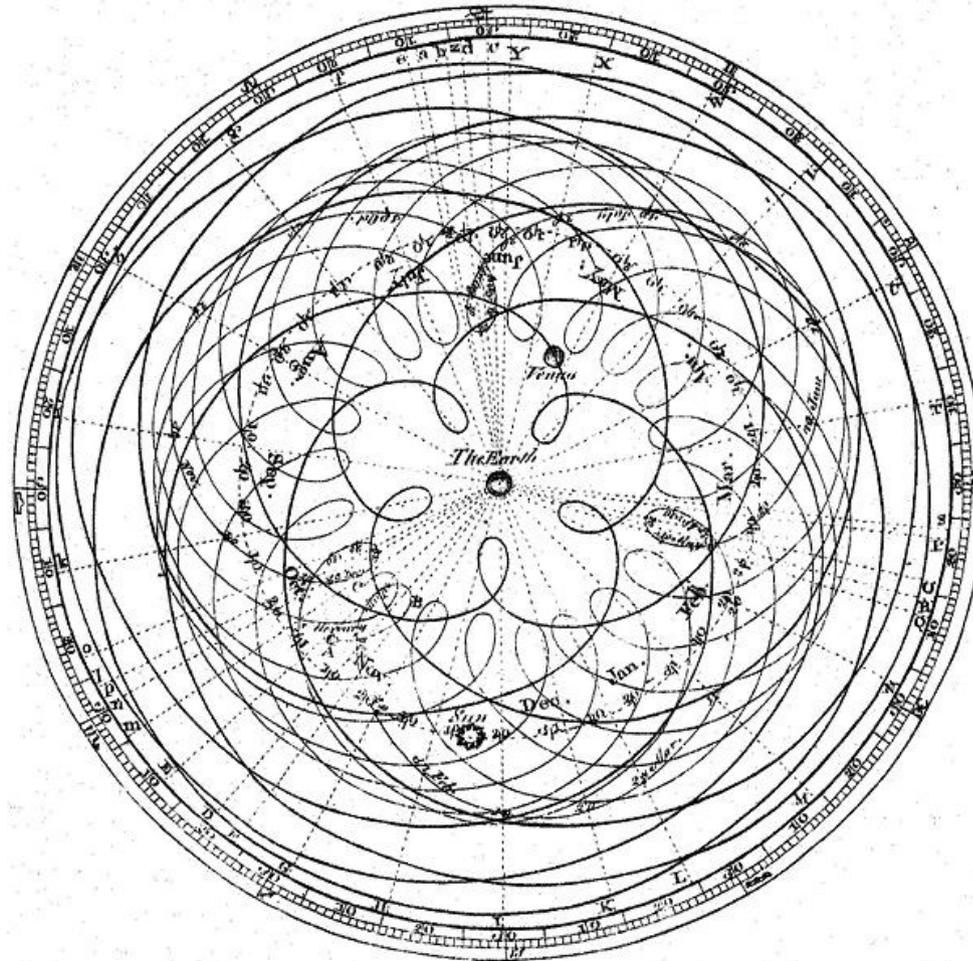


‘Lamarckian’ mechanisms in darwinian evolution

**Eva Jablonka
Marion J. Lamb
Eytan Avital**

14 mai 2013

Epigénétique et Variabilités Génomiques



« Epi... »

Épicycles : décrits dès le 3ème siècle avant JC, éléments du système géocentrique de Ptolémée (2ème siècle après JC)

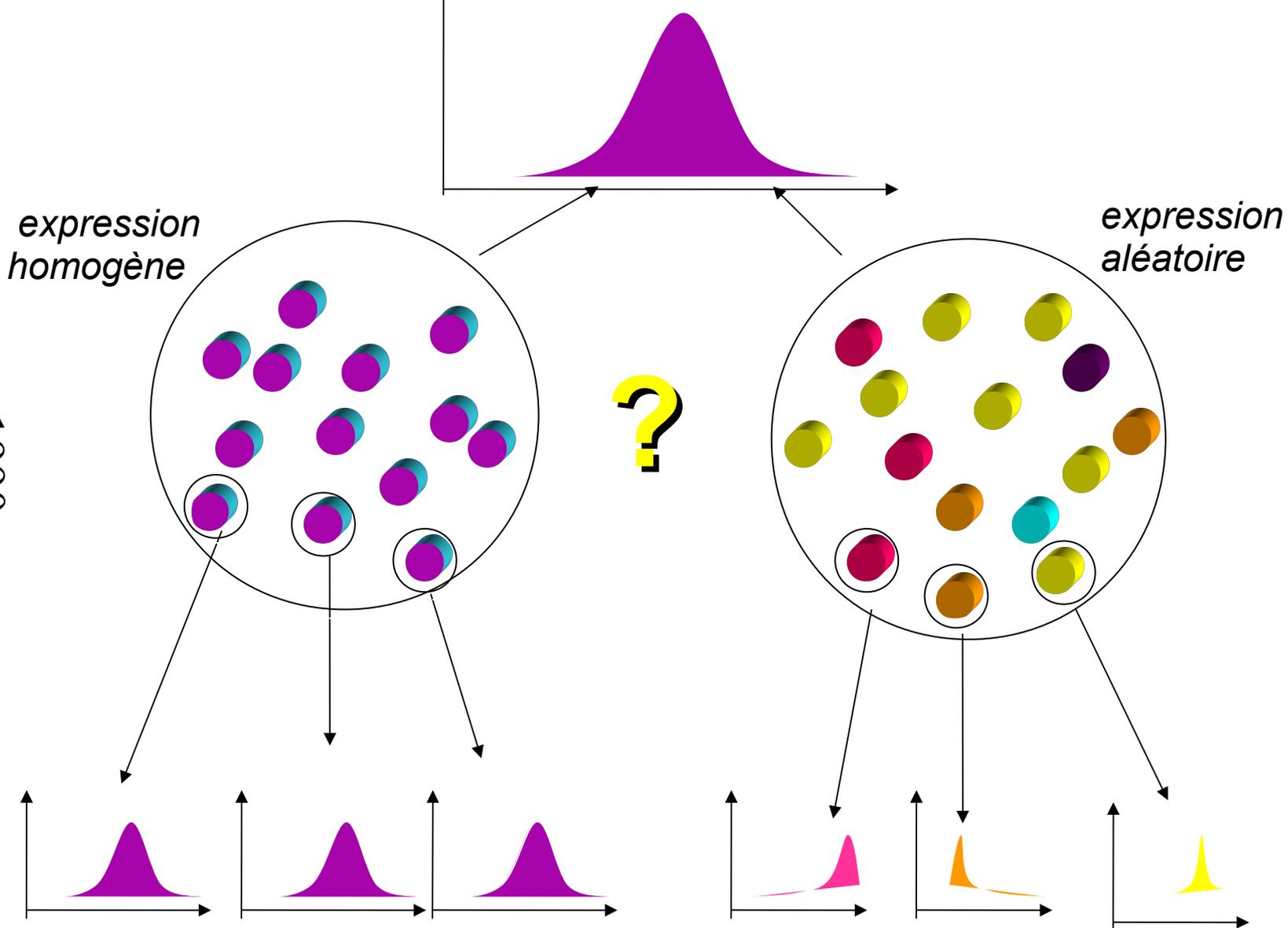
Abandonnés à la fin de la révolution copernicienne...

Une expression *stochastique* des gènes ?

“The mechanism causing cats to beget cats and fish to beget fish is hardwired in the genomic DNA (...) ”
(*Davidson et al, 2002*)

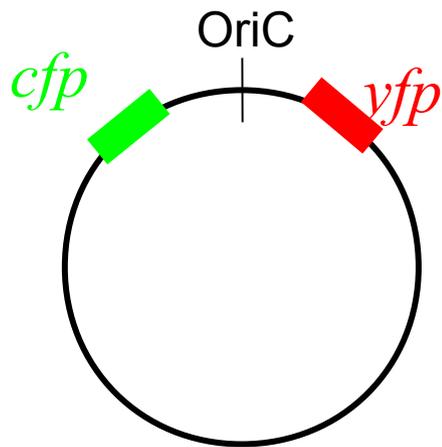
“The large number of cell states that are present in the lifetime of an organism and the reproducibility with which they are generated indicates the existence not just of programmes but also of mechanisms that ensure their reliable execution”
(*Arias & Hayward, 2006*)

Expérimentation sur une population de cellules

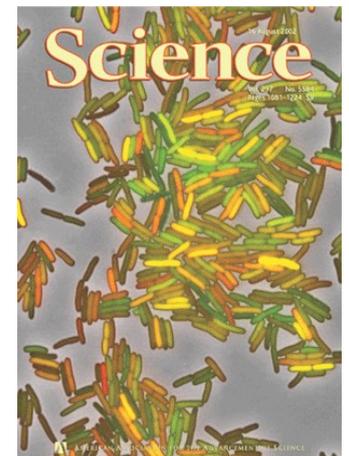
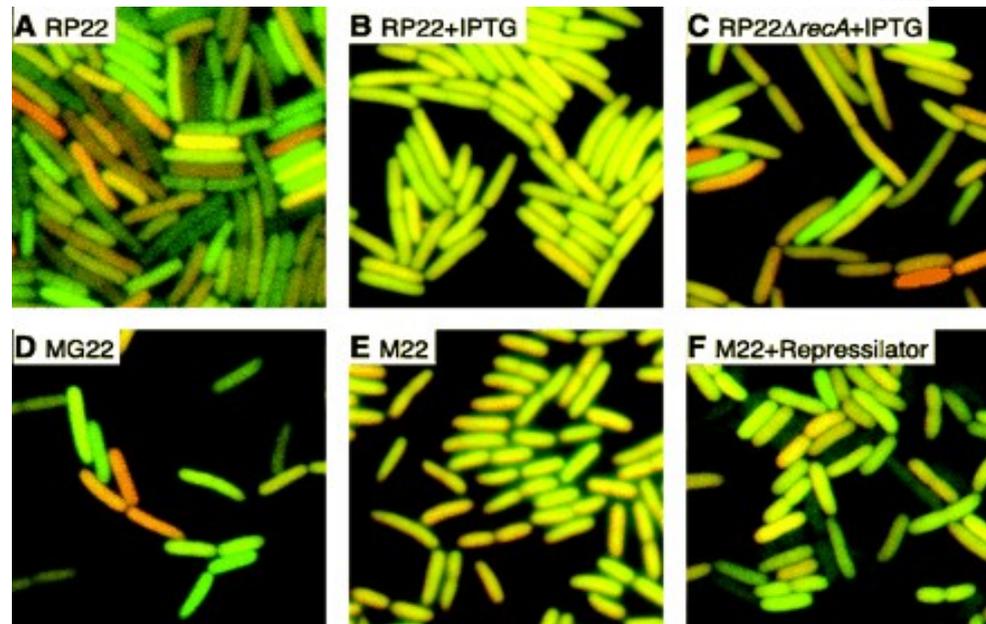
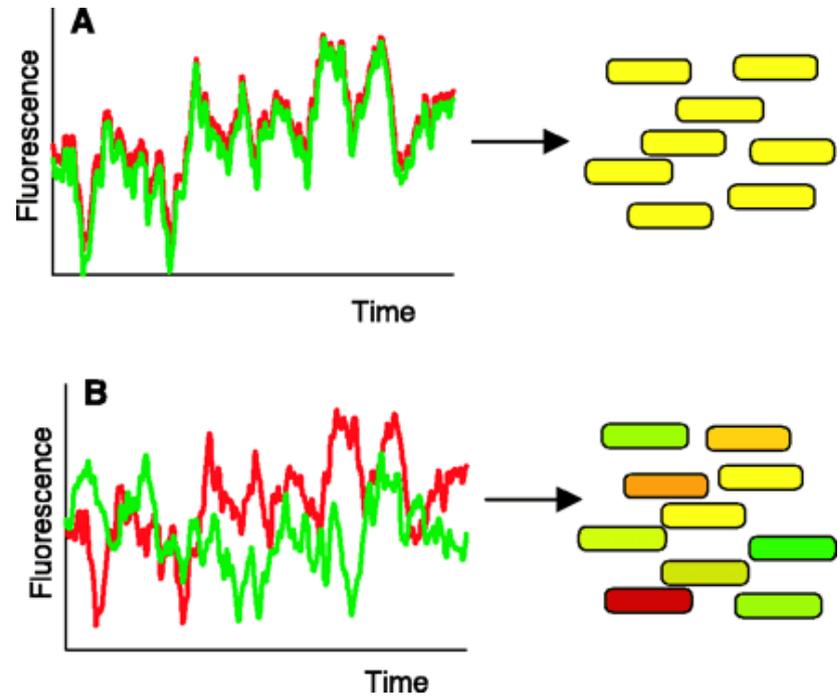


14 mai 2013

Epigénétique et Variabilités Génomiques

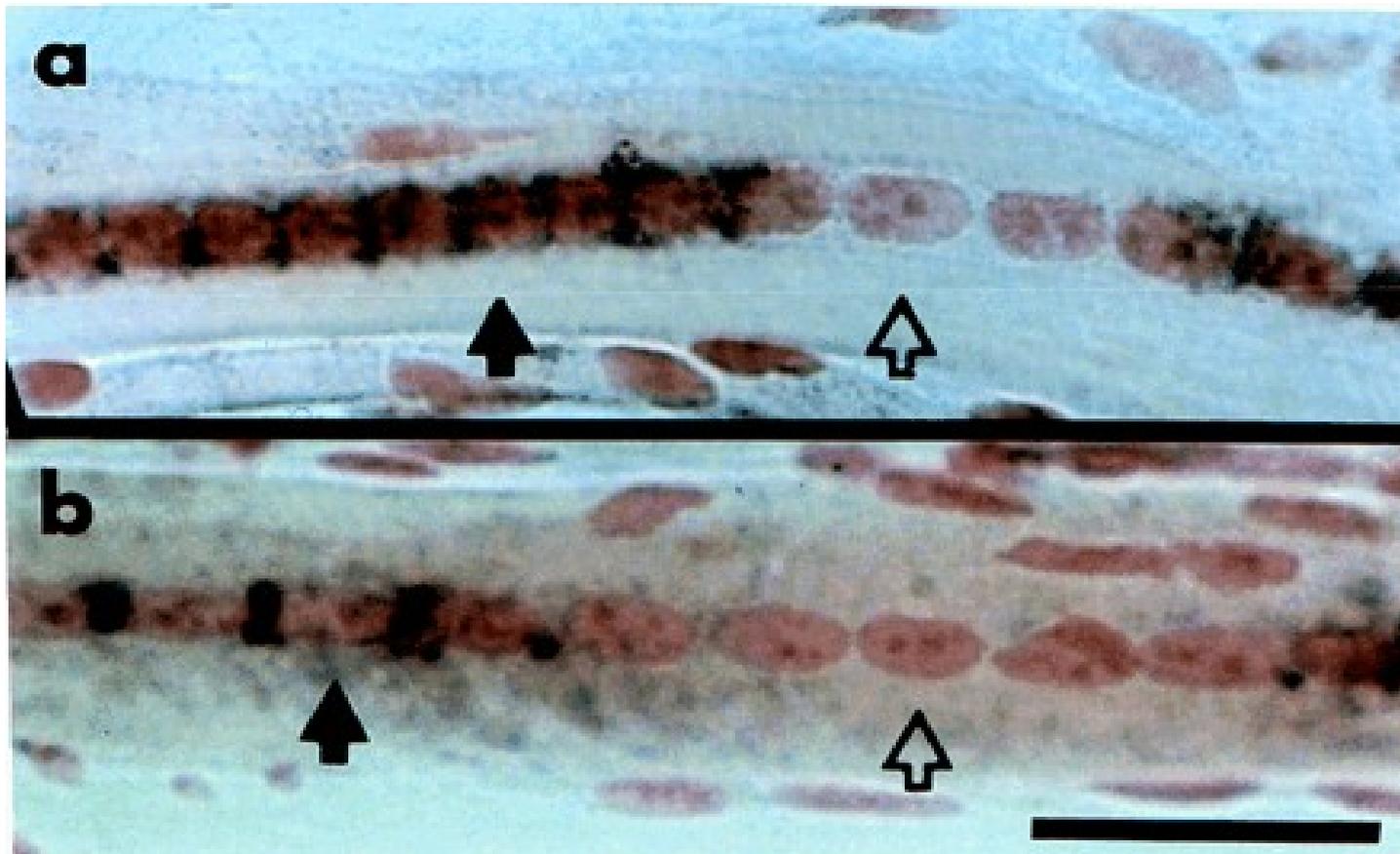


Elowitz et al, 2002



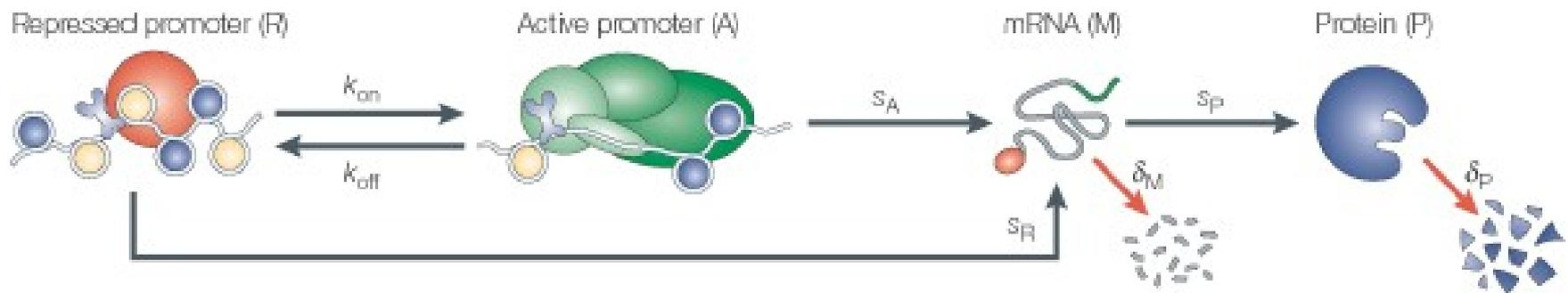
14 mai 2013

Epigénétique et Variabilités Génomiques



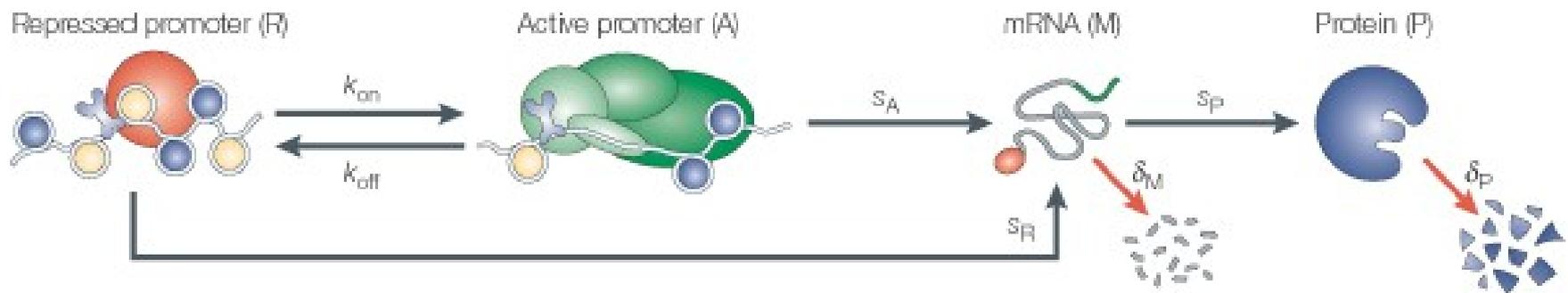
Newland *et al*, 1998

Expression différentielle du gène α -SA au sein d'une fibre musculaire multinucléée (souris)



L 'expression génétique est un processus multiétapes

Kaern et al, 2005



Kaern et al, 2005

L'expression génétique est un processus multiétapes

Les structures nucléaires sont très dynamiques

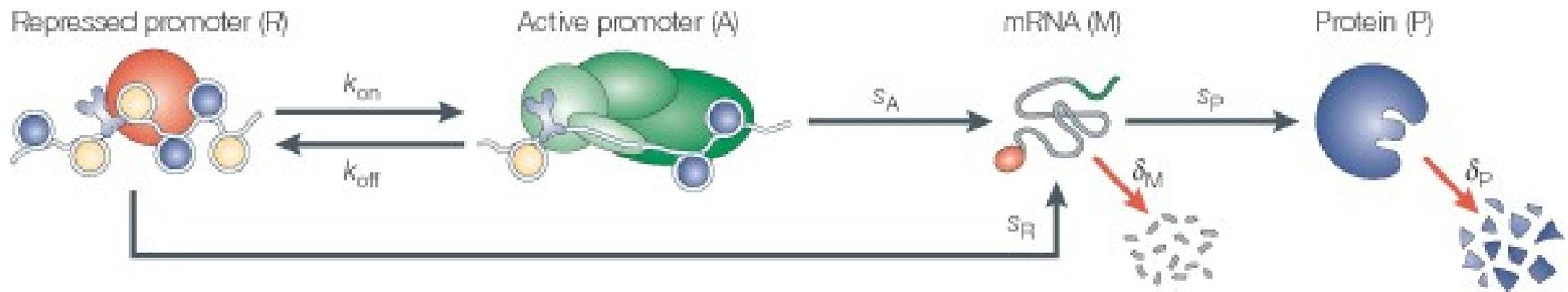
Exemples : Chromatine \longleftrightarrow H1 : 1 minute

Activateurs de la transcription (parfois <1 /cellule !)

Complexe d'initiation de la transcription

"Usines" à transcription : \nearrow Probabilité transcription

Localisation : 2500 sites *seulement*



Kaern et al, 2005

L'expression génétique est un processus multiétapes

Les structures nucléaires sont très dynamiques

Exemples : Chromatine \longleftrightarrow H1 : 1 minute

Activateurs de la transcription (parfois <1 /cellule !)

Complexe d'initiation de la transcription

"Usines" à transcription : \nearrow Probabilité transcription

Localisation : 2500 sites *seulement*

La disponibilité des molécules régulatrices est fluctuante

80% des protéines : -100 copies/cellule

Exemples : PBP2 (division cellulaire) 20 copies/cellule

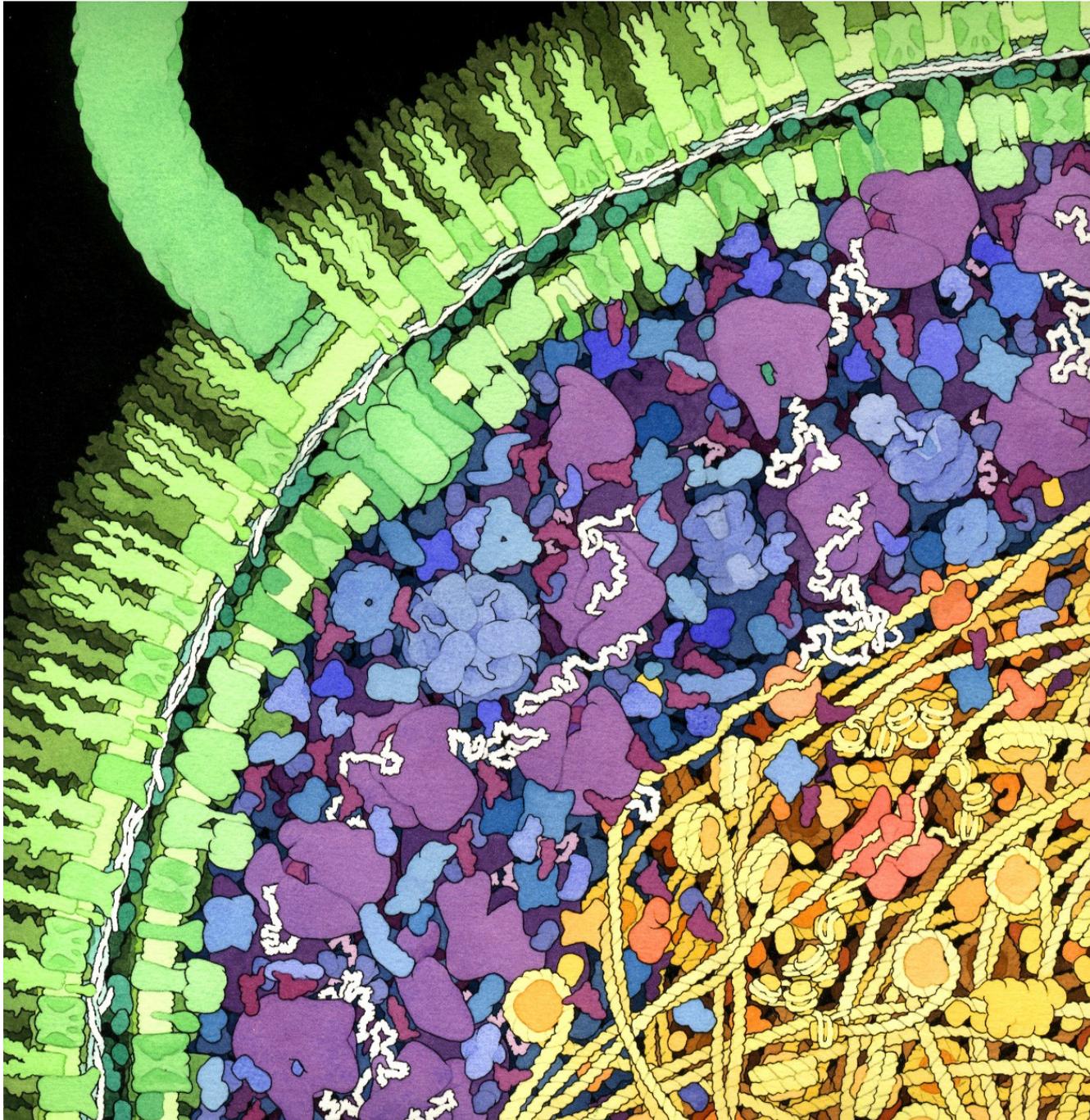
Primase dnaG (initiation réplication) 50 à 100

Deoxyribopyrimidine photolyase (réparation ADN): 10 à 20

\longrightarrow + **Effet d'échantillonnage**

14 mai 2013

Epigénétique et Variabilités Génomiques



**Encombrement
macromoléculaire**

Goodsell, 1993

14 mai 2013

Epigénétique et Variabilités Génomiques

Influence of Stochastic Gene Expression on the Cell Survival Rheostat after Traumatic Brain Injury

Daniel R. Rojo¹, Donald S. Prough¹, Michael T. Falduto², Deborah R. Boone¹, Maria-Adelaide Micci¹, Kristen M. Kahrig¹, Jeanna M. Crookshanks¹, Arnaldo Jimenez³, Tatsuo Uchida¹, Jeremy C. Cowart¹, Bridget E. Hawkins¹, Marcela Avila¹, Douglas S. DeWitt¹, Helen L. Hellmich^{1*}

¹ Department of Anesthesiology, University of Texas Medical Branch, Galveston, Texas, United States of America, ² GenUs BioSystems, Northbrook, Illinois, United States of America, ³ VeL- Lab Research, Missouri City, Texas, United States of America

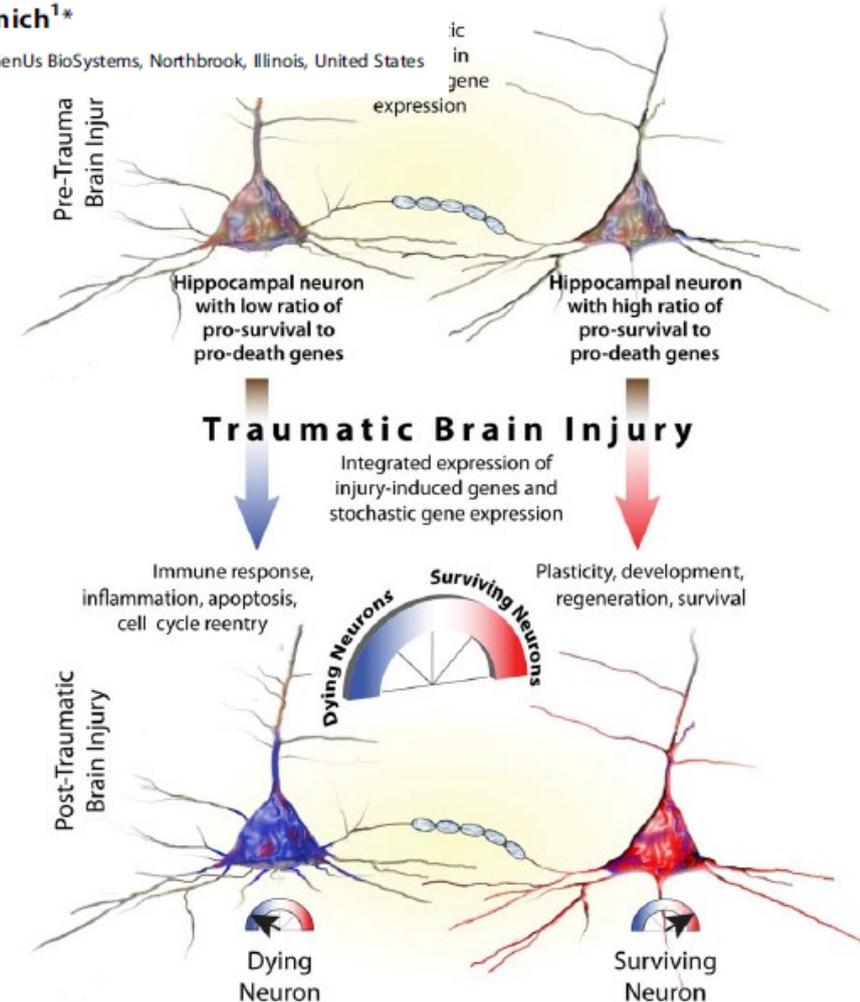
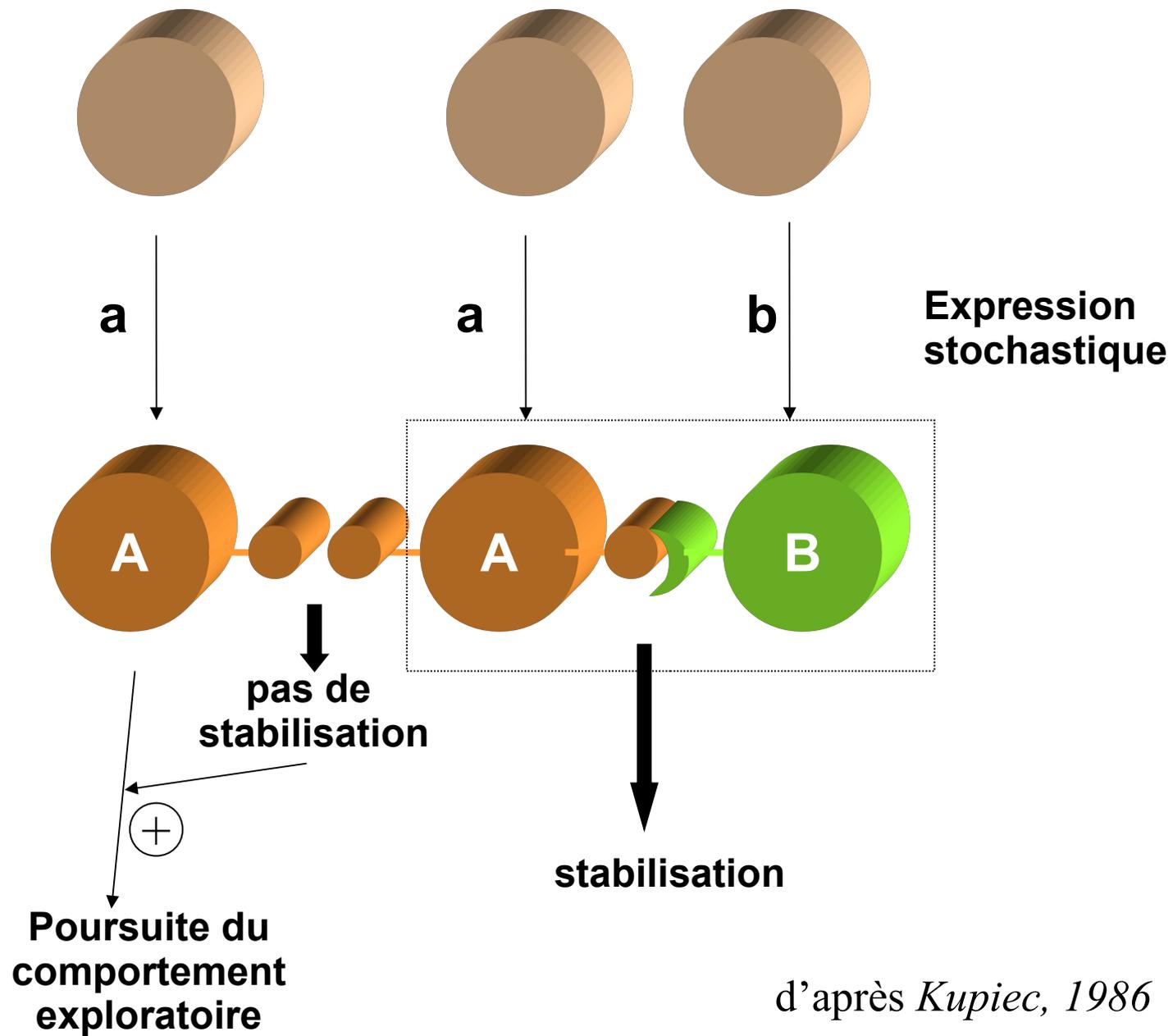
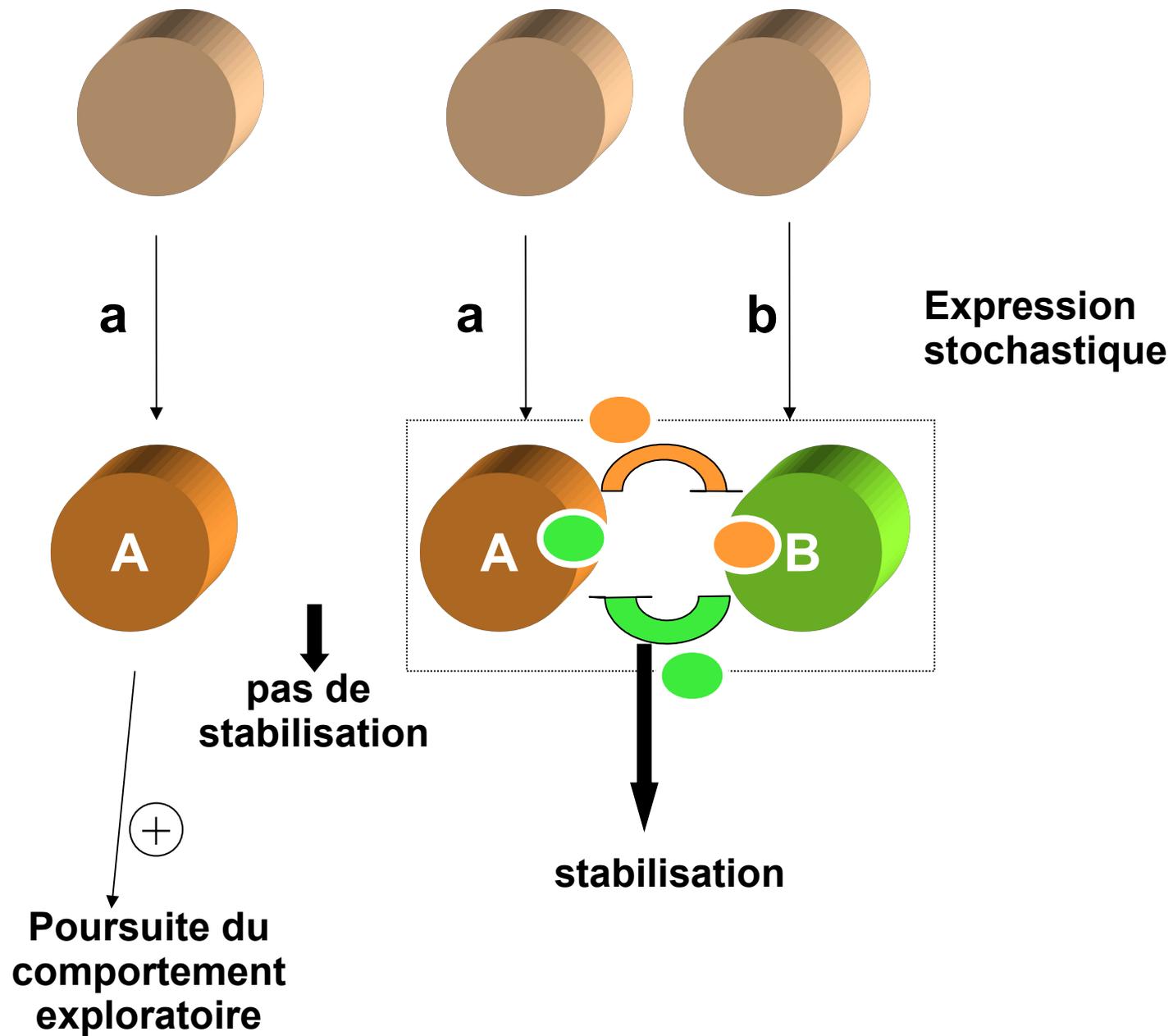
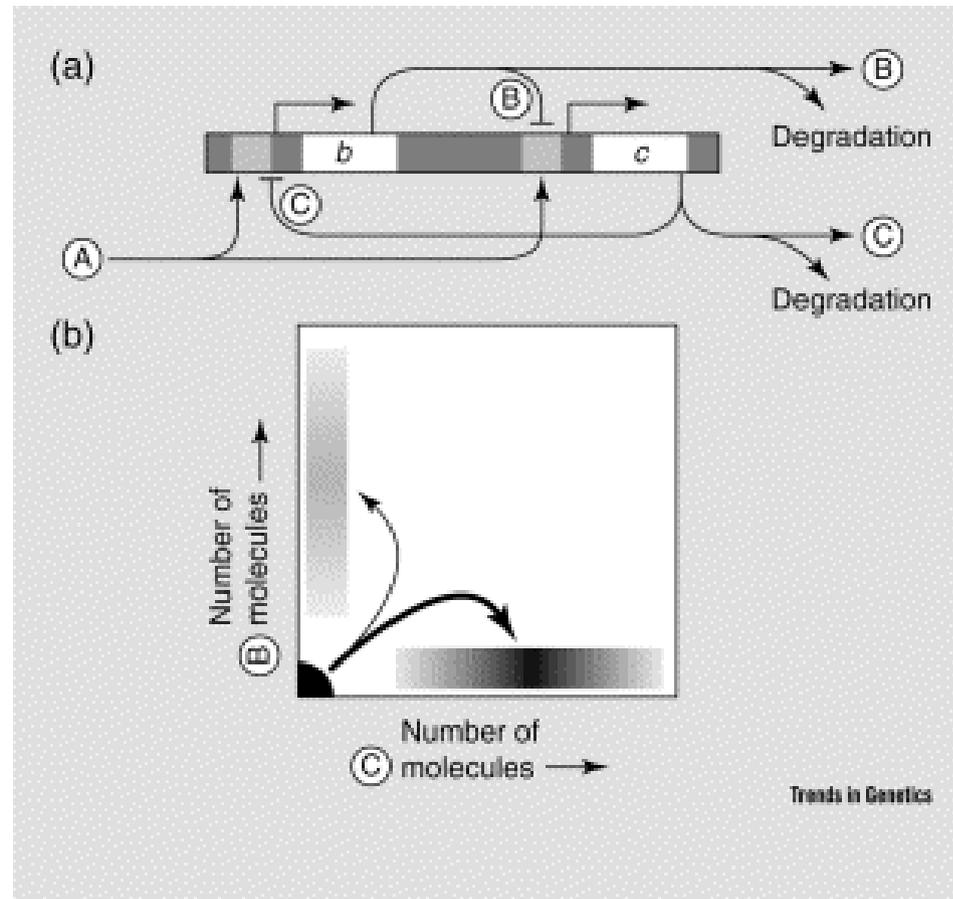


Figure 7. Rheostat model of neuronal survival after traumatic brain injury. doi:10.1371/journal.pone.0023111.g007



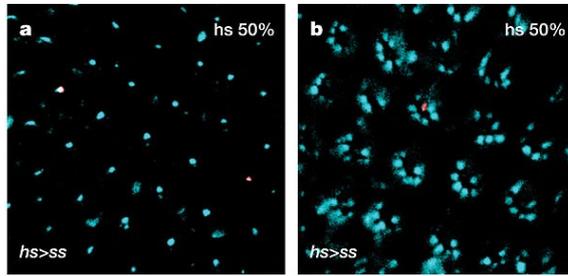


Fonctionnement déterministe et résultat aléatoire

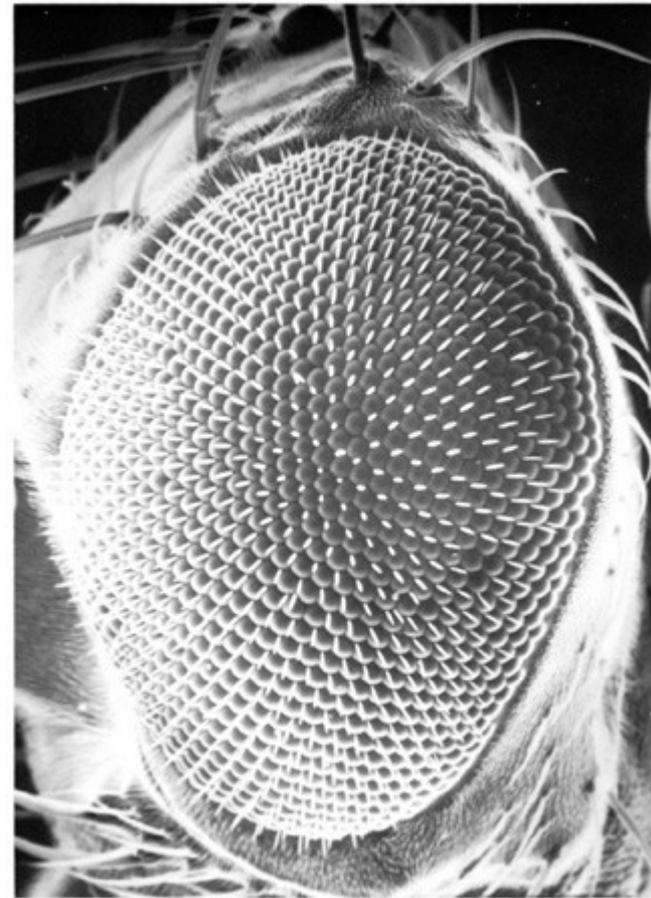
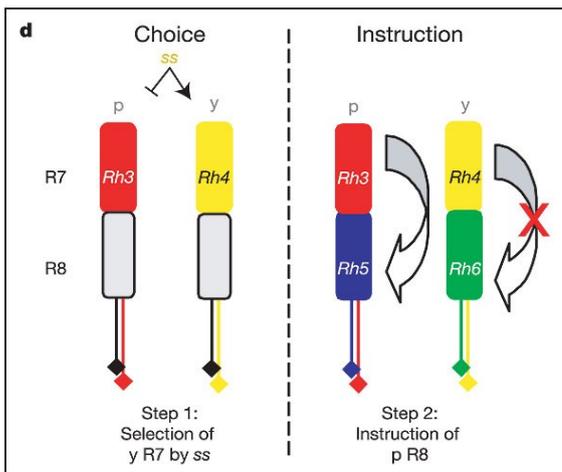
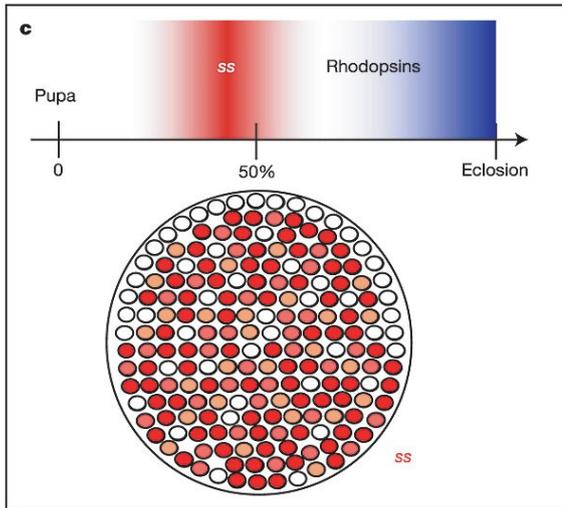


Arkin & McAdams, 1999

Expression stochastique de *spineless* et différenciation



Rh3 Rh4



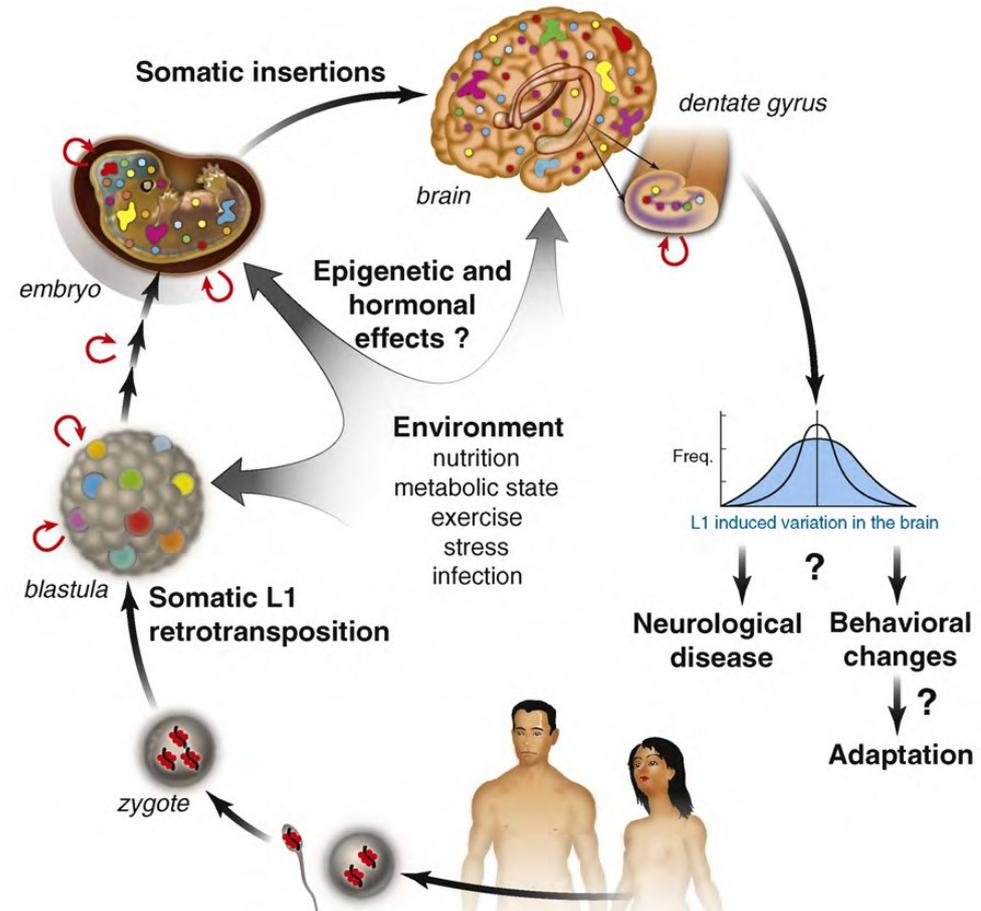
Wernet et al, 2006

14 mai 2013

Épigénétique et Variabilités Génomiques

Un génome mosaïque

LINE-1 (L1) elements are retrotransposons that insert extra copies of themselves throughout the genome using a 'copy and paste' mechanism. L1s comprise nearly ~20% of the human genome and are able to influence chromosome integrity and gene expression upon reinsertion. Recent studies show that L1 elements are active and 'jumping' during neuronal differentiation. New somatic L1 insertions could generate 'genomic plasticity' in neurons by causing variation in genomic DNA sequences and by altering the transcriptome of individual cells. Thus, L1-induced variation could affect neuronal plasticity and behavior. We discuss potential consequences of L1-induced neuronal diversity and propose that a mechanism for generating diversity in the brain could broaden the spectrum of behavioral phenotypes that can originate from any single genome.



Singer *et al*, 2010

Copy Number Variation across European Populations

Wanting Chen¹, Caroline Hayward², Alan F. Wright², Andrew A. Hicks³, Veronique Vitart², Sara Knott⁴, Sarah H. Wild⁵, Peter P. Pramstaller^{3,6,7}, James F. Wilson⁵, Igor Rudan^{5,8}, David J. Porteous^{1*}

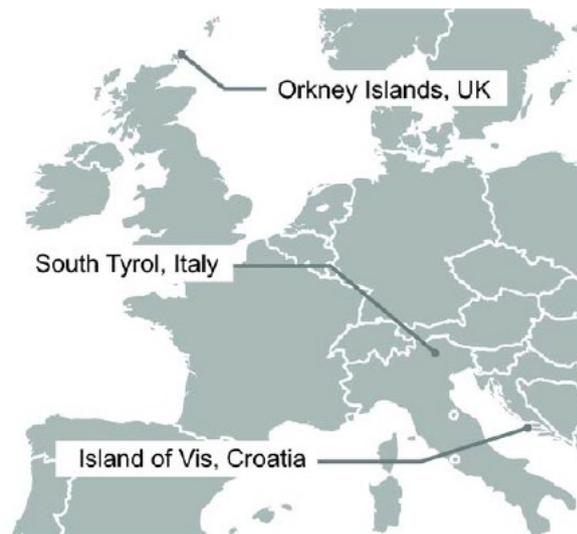


Figure 1. Geographic distribution of study samples.

Table 1. Characteristics of Copy Number Variants (CNVs) in Dalmatian, Orcadian and South Tyrolean populations.

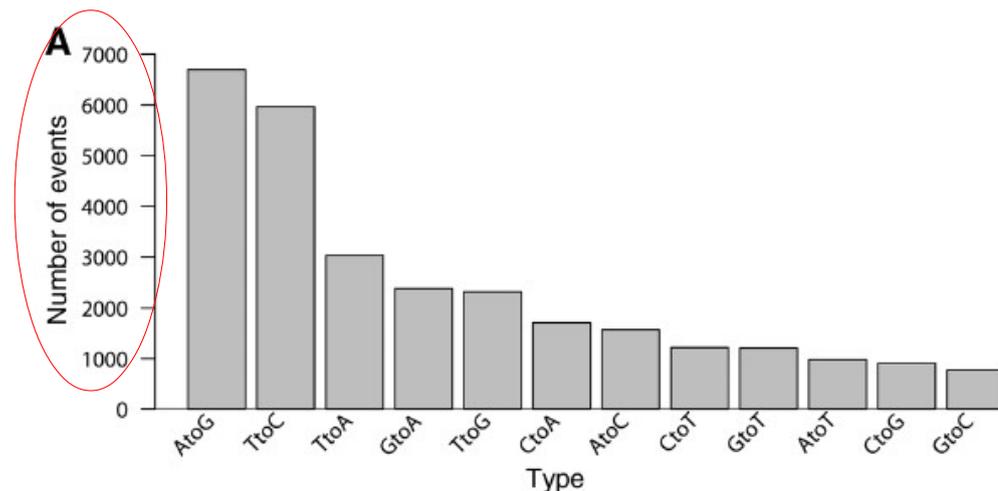
Population	Sample size	CNV carriers (percentage of carriers in population)	Number of CNVs	CNVs per person	Amplifications	Deletions	CNV mean length (kb)
Vis	965	702 (72.7%)	1384	1.43	803	581	216
Orkney	691	367 (53.1%)	630	0.91	324	306	192.6
South Tyrol	1133	895 (79.0%)	2002	1.77	1033	969	201.6
Combined	2789	1964(70.4%)	4016	1.44	2160	1856	205.1

doi:10.1371/journal.pone.0023087.t001

Widespread RNA and DNA Sequence Differences in the Human Transcriptome

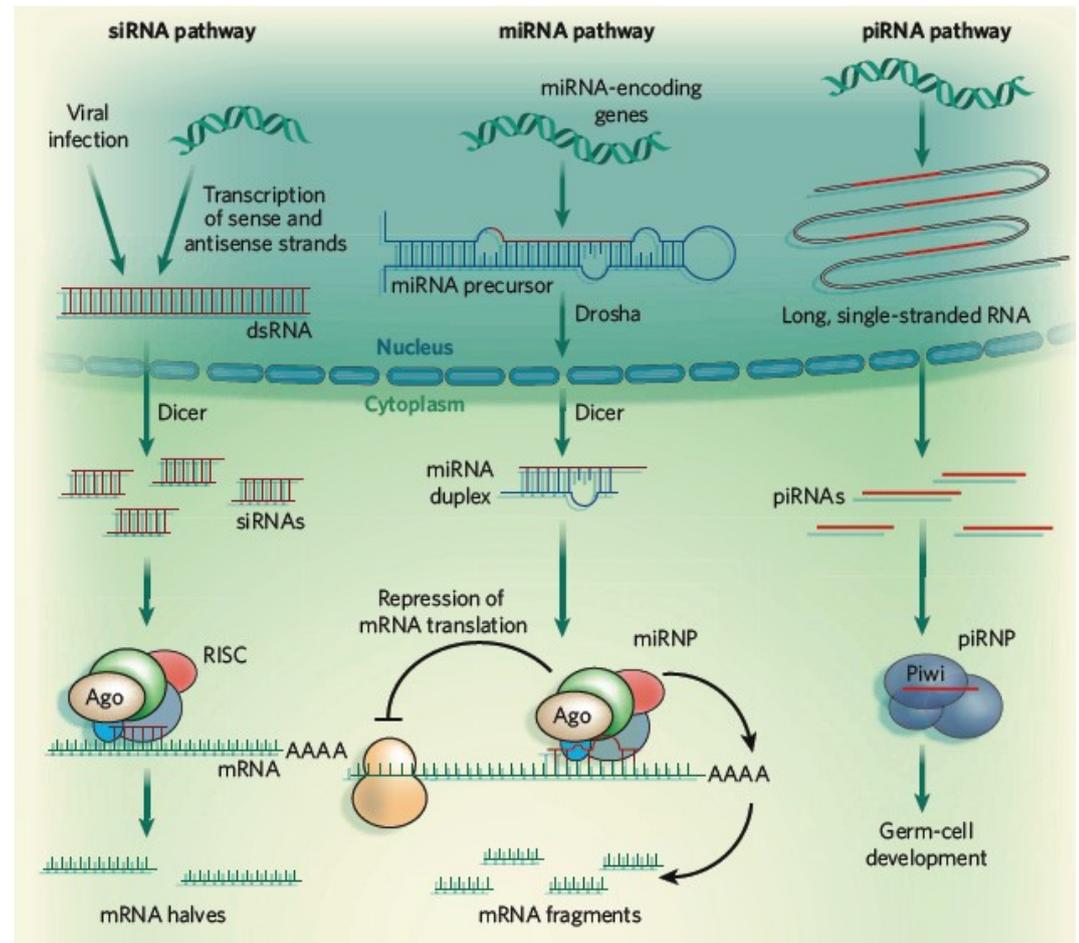
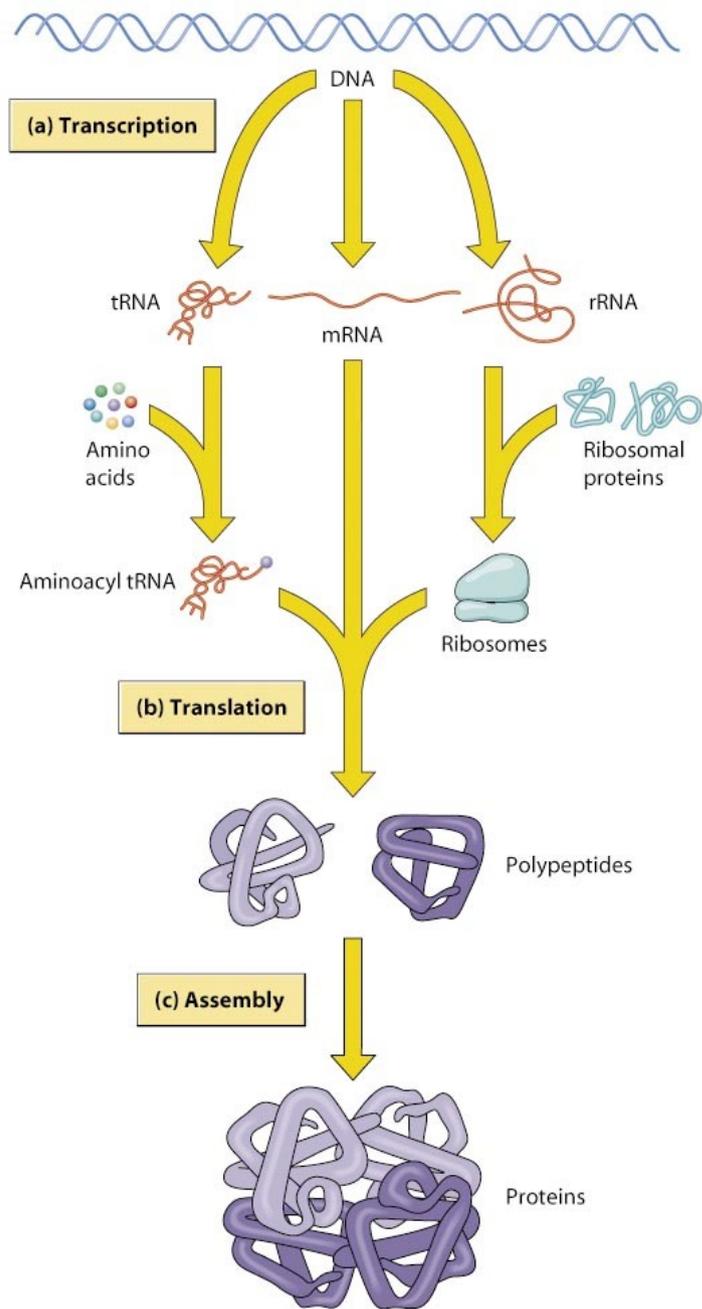
Mingyao Li,^{1*} Isabel X. Wang,^{2*} Yun Li,^{3,4} Alan Bruzel,² Allison L. Richards,⁵ Jonathan M. Toung,⁶ Vivian G. Cheung^{2,7,8†}

The transmission of information from DNA to RNA is a critical process. We compared RNA sequences from human B cells of 27 individuals to the corresponding DNA sequences from the same individuals and uncovered more than 10,000 exonic sites where the RNA sequences do not match that of the DNA. All 12 possible categories of discordances were observed. These differences were nonrandom as many sites were found in multiple individuals and in different cell types, including primary skin cells and brain tissues. Using mass spectrometry, we detected peptides that are translated from the discordant RNA sequences and thus do not correspond exactly to the DNA sequences. These widespread RNA-DNA differences in the human transcriptome provide a yet unexplored aspect of genome variation.



Mais... →

« Evidence of altered DNA stirs debate » *Nature News*, 2011 → ?



Copyright © 2009 Pearson Education, Inc.

14 mai 2013

Épigénétique et Variabilités Génomiques

RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription

Philipp Kapranov,¹ Jill Cheng,¹ Sujit Dike,¹ David A. Nix,¹ Radharani Duttagupta,¹ Aaron T. Willingham,¹ Peter F. Stadler,² Jana Hertel,² Jörg Hackermüller,³ Ivo L. Hofacker,⁴ Ian Bell,¹ Evelyn Cheung,¹ Jorg Drenkow,¹ Erica Dumais,¹ Sandeep Patel,¹ Gregg Helt,¹ Madhavan Ganesh,¹ Srinka Ghosh,¹ Antonio Piccolboni,¹ Victor Sementchenko,¹ Hari Tammana,¹ Thomas R. Gingeras^{1*}

2007

OPEN ACCESS Freely available online

PLOS BIOLOGY

Most “Dark Matter” Transcripts Are Associated With Known Genes

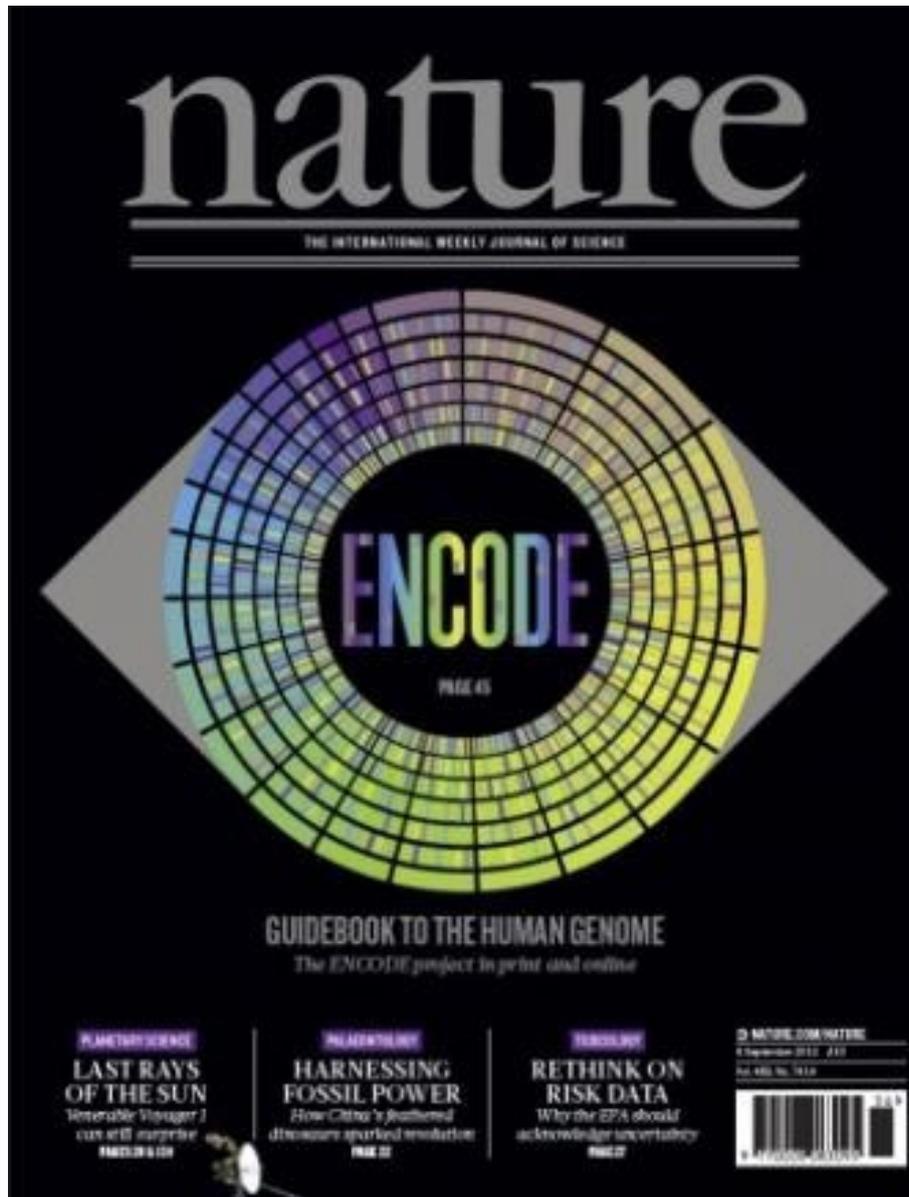
Harm van Bakel¹, Corey Nislow^{1,2}, Benjamin J. Blencowe^{1,2}, Timothy R. Hughes^{1,2*}

2010

The Reality of Pervasive Transcription

Michael B. Clark¹, Paulo P. Amaral^{1,9}, Felix J. Schlesinger^{2,9}, Marcel E. Dinger¹, Ryan J. Taft¹, John L. Rinn³, Chris P. Ponting⁴, Peter F. Stadler⁵, Kevin V. Morris⁶, Antonin Morillon⁷, Joel S. Rozowsky⁸, Mark B. Gerstein⁸, Claes Wahlestedt⁹, Yoshihide Hayashizaki¹⁰, Piero Carninci¹⁰, Thomas R. Gingeras^{2*}, John S. Mattick^{1*}

2011...



14 mai 2013

Epigénétique et Variabilités Génomiques

Merci de votre attention !

14 mai 2013

Epigénétique et Variabilités Génomiques

14 mai 2013

Epigénétique et Variabilités Génomiques